THE ACOUSTICAL FOUNDATIONS OF BEL CANTO

by

Cheruvathur Uthup

Submitted to the faculty of the
Jacobs School of Music in partial fulfillment
of the requirements for the degree,
Doctor of Music
Indiana University
December 2016

Accepted by the faculty of the
Indiana University Jacobs School of Music,
in partial fulfillment of the requirements for the degree
Doctor of Music

Doctoral Committee

____
Alice Hopper, Chair & Research Director

____
Costanza Cuccaro

____
Patricia Havranek

____
Peter Miksza

November 17, 2016

In loving memory of my grandfather, whose voice still rings in my heart

Acknowledgements

I would first like to express my heartfelt gratitude to Alice Hopper, my research director, chair, and teacher of 15 years. Her support, encouragement, and guidance have been such a blessing to me. I am truly indebted to her for all she has done for me. I would also like to thank Professors Patricia Havranek, Costanza Cuccaro, and Peter Miksza for graciously agreeing to serve on my dissertation committee. Their wisdom, insight, and guidance have been extremely instrumental in this project.

My family has always been a constant pillar of support to me, especially throughout the completion of this project. There are not enough words to fully express the extent of my gratitude to them. I would like to thank my wonderful father and mother, Simon and Jasmine Zachariah, for weaving music into my life and encouraging me to pursue my dreams; my father-in-law and mother-in-law, Rev. K.E. Easaw and Mercy Easaw, for their continuous love and prayers; my sister-in-law, Anugraha Easaw, for her words of encouragement. I would also like to thank my loving wife, Gifty Easow, for her unwavering devotion and constant support.

Last but most importantly, I want to thank God for his bountiful blessings and giving me the sufficient grace and strength to complete this project. May this work be used for the glory and extension of His kingdom.

Abstract

In the area of voice instruction, tradition and science are often viewed as being on opposite ends of the spectrum. Pitting science and tradition against each other in a mutually exclusive fashion frequently results in unnecessary confusion, antagonism, and misunderstandings among both students and teachers of singing. However, the traditional teachings of *bel canto* are not at odds with modern science. In fact, much of the traditional precepts have support in the modern scientific literature. Thus, the purpose of this project is to bring to light the underlying scientific principles present within the traditional teachings of *bel canto*.

This analysis is done in three stages. First, a brief definition and history of the *bel canto* tradition is discussed in which the reader can appreciate the historical context that gave rise to the principles and practices with which the term is often associated. Next a scientific groundwork is established, focusing on the essential acoustic, physiologic, and perceptual processes related to the art of singing. Finally, the precepts and teachings of the *bel canto* era are examined and synthesized with the previously established scientific groundwork. Through these three stages, the acoustical foundations of *bel canto* are brought to light.

Table of Contents

List of Figures

List of Abbreviations

CT:	cricothyroid
IA:	interarytenoid
LCA:	lateral cricoarytenoid
PCA:	posterior cricoarytenoid
TA:	thyroarytenoid

Chapter 1: INTRODUCTION AND PURPOSE

The *bel canto* style of singing, which essentially translates to "beautiful singing," has been around from the middle of the 16^{th} century to the beginning of the 19^{th} century. The advent of monody and Giulio Caccini's treatise, *Le Nuove Musiche*, laid the foundation for the germination and blossoming of the *golden age of bel canto*. The art of *bel canto* was primarily passed down from teacher to student by oral tradition, perhaps in order to guard trade secrets. As a result, our knowledge of *bel canto* technique is fairly limited. Although there are several recurring concepts and themes in the available literature which provide an approximation of the technique, it is still difficult to reconstruct the precise methods and practices specific to the era. Furthermore, the abstract and artistically-oriented instruction that appears in many *bel canto* publications (e.g., Lamperti's instruction in *Vocal Wisdom*) also make it difficult to grasp, especially in this age of increasing scientific awareness.

Art and science are both integral aspects of beautiful singing. Yet, the schism between science and art in the area of voice instruction has been present ever since the invention of Manuel García's laryngoscope. In fact, this schism has grown wider throughout the years as a result of new scientific discoveries and theories being made each year. Although several vocal pedagogues, most prominently William Vennard and Richard Miller, began to bridge this gap in the mid-20th century by bringing scientific awareness into the realm of artistic instruction, there still remains a prominent gap in vocal instruction.

Although understanding the scientific principles underlying *bel canto* technique may not be entirely necessary to learning it, there are several benefits in doing so. First, this understanding will allow the student to be confident that the specific ideas and exercises being taught are grounded in science. Second, this understanding is integral to teachers when students question the rationale of the techniques. Finally, and perhaps most importantly, a strong foundation in human physiology as it relates to *bel canto* principles will help demystify many of the erroneous beliefs and practices that lead young vocalists astray.

This project aims to shed light on the scientific principles underlying *bel canto* technique by combining the principles from the available literature on *bel canto* technique together with the current scientific literature. After providing a brief history and principles of the *bel canto* style of singing, the aspects of sound production, physiology of singing, and perceptual aesthetics will first be analyzed through a scientific lens and subsequently through the lens of *bel canto* technique. This process will allow for a basic understanding of the underlying mechanisms of each aspect as well as their significance to *bel canto* principles.

In order to provide a clear understanding of the history and principles of the *bel canto* style, the historical treatises and teachings of the period are analyzed as primary sources. Furthermore, to supplement this analysis and to obtain a more comprehensive understanding, information from secondary sources such as books, articles, dissertations, and manuals on *bel canto* technique are incorporated into the analysis. Subsequently, a thorough understanding of the scientific principles of sound production, physiology of singing, and perceptual aesthetics are supported by secondary sources from multiple disciplines such as acoustics, medicine, speech and hearing, and psychology.

This study begins with a review of the history and development of the *bel canto* style of singing. The main schools of singing are discussed in detail along with their methods and philosophies. After the historical investigation into the development of the *bel canto* style, a scientific groundwork is laid with which to appreciate the acoustic foundations of *bel canto* principles. The scientific groundwork focuses on topics such as the basic properties of sound, the principles and physiology of singing, as well as sensation and aesthetics. The main goal of these chapters is to shed light on the underlying phenomena occurring in basic sound production and perception. Once the scientific groundwork has been laid, the study aims to uncover the scientific processes inherent within the principles of *bel canto*.

Themes that emphasized in this study include theories of voice production, such as the source-filter theory and the non-linear source filter theory, models of sound perception and emotion, as well as principles of the basic physics of sound. By understanding how these themes are connected to the core principles of *bel canto*, singers can gain a better understanding of the acoustical foundations of *bel canto* singing.

Chapter 2: THE ART OF BEL CANTO

Definition of Terms

The term *bel canto*, despite its simple and straightforward translation, is frequently the epicenter of heated debates and misunderstandings among singers, teachers of singing, musicologists, and other scholars of music. This stems from the fact that the term often evokes a prismatic array of meaning. Like many of the terms surrounding vocal instruction, there are multiple perspectives on the term *bel canto*. Some scholars maintain that it refers strictly to a style of composition that flourished particularly in 18th and 19th century Italian opera, in which embellishment and ornamentation were of primary importance (Celletti, 1991; Franca, 1957). Others use the term to denote exclusively the vocal techniques and practices that were employed to attain the virtuosic and highly embellished style of the period (Duey, 1951; Stark, 1999). The term is sometimes even used by many in a much broader sense to refer to all vocal performance between 1600 and 1850.

Since *bel canto* is a frequently misinterpreted term due to its multiple shades of meaning, it is necessary to specify the intended meaning, especially in an in depth discussion such as this. Robert Toft (2013) explains that the term, which did not emerge until the latter half of the 19th century, was used by musicians of the time, who thought they were witnessing the dissipation of vocal music, simply to describe the loss of the old Italian vocal practices that were prevalent in the 18th to the early 19th centuries. For the purposes of this document, the term *bel canto* will hereafter refer to the above-mentioned Italian techniques and practices associated with the distinct vocal style of the period.

However, in order to fully understand these techniques and practices, it is first essential to understand the historical context in which they originated.

History

The emergence of solo singing as an art form happened as early as the end of the 16^{th} century. Although it was first concentrated in a few cities in northern Italy, a new class of solo singers gradually began to demonstrate vocal dexterity that surpassed that of the average choristers of the time. In Ferrara, a group of singers known as the *concerto delle donne* or "consort of ladies," which consisted of certain ladies of the court, became widely known for their technical and artistic virtuosity. They mainly performed madrigals, which were originally composed as part-songs, as solo songs with continuo. Although the group was formed in the court of Alfonso II d'Este of Ferrara and intended to be heard only by an exclusive audience, their style of singing was said to have greatly influenced the development of the seconda practica.

In the late 16^{th} century, when the composer, singer, and voice teacher Giulio Caccini (1551-1618) visited the court of Alfonso II and heard the *concerto delle donne*, he was so impressed and inspired by them that he established another consort of singers at Florence, sponsored by the Grand Duke Francesco de' Medici, to rival them (Stark, 1999, p. 194). With the formation of this group, the focus shifted from Ferrara to Florence. Caccini developed several novel ideas and techniques, which were at the time nothing short of revolutionary, in order to maximize the emotional impact of the text. His style of teaching also brought about a certain degree of finesse and nobility in his consort's singing. Since Caccini was first and foremost a singer and teacher of singing, he composed his music to best utilize his new techniques, which he compiled in his publication, *Le Nuove Musiche*.

Giulio Caccini (1602), in his introduction to his famous *Le Nuove Musiche*, described clearly and in great detail the aesthetics of his newly developed style. In his lengthy discussion of the various forms of ornamentation, he disdained the lavish and ignorant uses of ornaments and instead maintained that ornaments should be used to create a desirable affect. In his description of the conception of his style, which is largely based on affect, he borrows the words of Plato to defend his teachings: “music is naught but speech, with rhythm and tone coming after” (Caccini & Hitchcock, 1970, p. 44). Caccini made it clear that he believed affect and clear understanding of the text to be the most important aspect of music and therefore paramount to the form and the content of the piece.

According to Caccini, the music in the old style “offered no pleasure beyond that which pleasant sounds could give—solely to the sense of hearing, since they could not move the mind without the words being understood” (Caccini & Hitchcock, 1970, p. 44). Caccini’s new style of composition differed from the styles of his predecessors in that he prioritized affect so much so that he sacrificed the inner voices altogether to create a completely new texture: simply solo voice and continuo. He defended this new style by claiming that “a single voice has more power to delight and move than several voices together” (Caccini & Hitchcock, 1970). This new genre, which is now referred to as monody, laid the foundation for all solo vocal music to come.

Caccini’s school of singing in the early 17th century was the first of many golden ages of singing, and it established an Italian vocal tradition that would be passed on from generation to generation and eventually considered the wellspring of *bel canto* (Stark, 1999). Henry Pleasants (1966) explains that although there were several golden ages of singing, they all shared one unifying characteristic—they were all continuously dominated by voices trained in the Italian school of singing. The unceasing dominance of the Italian tradition is further highlighted by the fact that, whenever the various schools of singing are discussed in the historical literature, the Italian school is regarded as the model against which all other schools are measured.

Throughout the years there were several peaks in good singing dominated by clusters of great singers which are now referred to by scholars as golden ages. Henry Pleasants (1966) writes about four distinct golden ages of singing apart from the first one in the 17^{th} century during Caccini's time. In the eighteenth century, there was a golden age that extended from about 1720 to 1740 and another from 1770 to 1790. These two golden ages were primarily dominated by the castrati—male singers who were castrated at a young age in order to preserve the larynx of a boy while acquiring the physique, musculature, and lung capacity of an adult male. One factor that contributed to the rise of the castrati was the value of celibacy in a poor economy. It was not uncommon for parents of humble means to castrate their sons and send them to a chapel choir in order to ensure a steady living for them. Moreover, castration would prevent the creation of any additional children to feed.

Castrati were especially popular due to the fact that their vocal ability, in terms of range, power, flexibility, and unique timbre, usually surpassed those of female singers of the time. In fact, castrati were so popular that the singer and composer, Pier Francesco Tosi, who himself was a castrato, complained that "Italy hears no more such exquisite voices as in times past, particularly among the women" (Tosi & Galliard, 1968, p. 15). Another factor that helped to propagate the popularity of the castrato voice was the ancient belief that sexual ambiguity, particularly the hermaphrodite, represented the supernatural (Stark, 1999). In fact, many castrato roles portrayed mythological characters or gods. The castrato voice offered something new and extraordinary that was not heard of before, and this captivated audiences of the time. Although these golden ages produced numerous outstanding singers and teachers, one of the most influential was the great voice teacher Nicolò Porpora (1686-1768), who revived the teachings of Caccini and further developed them. Among his numerous pupils were two renowned castrati: Farinelli (1705-1782) and Caffarelli (1710-1783) as well as a young Franz Joseph Haydn (Pilotti, 2009).

Despite the immense popularity of the castrati in the 18th century, their popularity faded by the beginning of the 19th century. The decline of the castrati occurred due to several reasons. First, economic situations in Italy improved, which caused individuals to become increasingly reluctant to castrate their sons. Moreover, in the middle of the 18th century, Pope Benedict XIV shunned the act of castration, calling it "an unnatural crime, the victims of which are young boys, often through the complicity of their parents" (Barbier, 1996). Another major factor that hastened the decline of the operatic castrato was the growing popularity of *opera buffa*, or comic opera. This genre of opera called for natural singers to play the roles of common stock characters. Thus the mythological characters used in *opera seria*, who were often played by castrati, were quickly overshadowed by the common everyday characters of *opera buffa*, which called for natural singers. All these factors together contributed to the diminishing popularity of the castrato singer.

The third golden age manifested around 1825-1840 after the decline of the castrati. This era was particularly important because it marked the start of the age-long rift between tradition and science in *bel canto* instruction. The two prominent schools of singing that emerged during this time were the Garcia School, founded by Manuel Garcia I (1775-1832) and greatly developed by his son Manuel Garcia II (1805-1906), and the Lamperti School, headed by Francesco Lamperti (1811-1892) and later on by his son Giovanni Battista Lamperti (1839-1910). Although both schools produced many great singers, they had very different styles of teaching. The Lamperti School was a traditional school of voice that heavily emphasized sensations, tone colors, and breath control using long-established vocal terminology. On the other hand, the Garcia school, which was cutting-edge at the time, was known for its technical approach grounded in scientific principles. Much of the school's reputation was established by Manuel Garcia II, whose pioneering work brought new scientific awareness to the area of vocal instruction. His contributions to the study of the voice were so significant that his treatises are still considered essential to contemporary scholars of singing.

Despite the scientific reputation it received later due to the work of Garcia II, the Garcia school was originally founded by Manuel Garcia I, the father of Garcia II, who, although Spanish born, was a singer and teacher of the traditional Italian method. He was not only a great singer, but he was also Gioachino Rossini's favorite tenor. In fact, Rossini created several roles especially for him, including the role of Count Almaviva in *Il Barbiere di Siviglia* (Stark, 1999, p. 3). He passed down this traditional Italian method to his three children: Maria Malibran, Pauline Viardot, and Manual Garcia II. In fact, he valued the traditional Italian style of singing so much that he sent his son, Manuel Garcia II, to Naples to study with Giovanni Ansani, who was himself a pupil of the renowned voice teacher Nicolò Porpora (Pilotti, 2009). Although all three of his children became well known in the field of singing, it is because of the life work of Manuel Garcia II that the school became so well known and associated with its distinct scientifically-oriented style.

Manuel Garcia II was a pivotal figure in both singing and medicine, primarily due to his groundbreaking invention: the laryngoscope. This invention was not only significant to the field of singing, but also to medicine in general. Before the invention of the laryngoscope, there was no other means of observing the vocal mechanism in action. Thus, his invention was a major scientific landmark. Garcia II was trained in the old Italian tradition, as mentioned earlier, and began singing professionally at a very early age. However, by the age of 24 he stopped performing due to vocal damage. Stark (1999) attributes his vocal injury to the strain of singing during his pubertal voice change and to singing leading roles at too young an age. After his injury, he became employed in military hospitals, during which time he witnessed several injuries of the head and neck and became intrigued with studying the anatomy and physiology of the human larynx. Although he invented the laryngoscope much later in his life, these experiences probably laid the foundation for his landmark achievement. After the death of his father in 1832, Garcia II took up the training of his two sisters in his place. Thus the success of their singing careers is at least in part attributed to his teaching. Three years later he was appointed to the position of Professor of Singing at the Paris Conservatoire, where he would publish his landmark treatise, *École de Garcia: Traité complet de l'art du chant*. The treatise was comprised of two sections that were published six years apart. The first part, published in 1841, was primarily focused on vocal technique, while the second part, which was added in 1847, mainly detailed the stylistic traditions and practices of the day. The treatise underwent several editions and revisions, including a two-part English translation of the 1872 version edited by his grandson, Albert Garcia.

Garcia's invention of the laryngoscope, a device that would later revolutionize the field of medicine, was his most valuable contribution by far. The following is an excerpt in Garcia's own words from his paper to the International Medical Congress of 1881 recounting the manner in which he received the inspiration for his breakthrough invention:

> One September day in 1854, I was strolling in the Palais Royal, preoccupied with the ever-recurring wish, so often repressed as unrealizable, when suddenly I saw the two mirrors of the laryngoscope in their respective positions, as if actually present before my eyes. I went straight to Charriere, the surgical instrument maker, and asking if he happened to possess a small mirror with a big handle, was informed that he had a little dentist's mirror which had been one of the failures of the London Exhibition in 1851. I bought it for six francs. Having obtained also a hand mirror, I returned home at once, very impatient to begin my experiments. I placed against the uvula the little mirror (which I had heated in warm water and carefully dried); then flashing upon its surface with the hand mirror a ray of sunshine, I saw at once, to my great joy, the glottis wide open before me, and so fully exposed that I could perceive a portion of the trachea. When my excitement had somewhat subsided, I began to examine what was passing before my eyes. The manner in which the glottis silently opened and shut and moved in the act of phonation, filled me with wonder. (MacCormac & Makins, 1881, n.p.)

Thus, by harnessing the illumination of the sun's rays by the use of mirrors, he was able to observe his own larynx through a process known as auto-laryngology. While some scholars argue that dental mirrors had previously been utilized by physicians to view parts of the larynx, it is widely acknowledged that it was Garcia who created the first laryngoscope in 1855. After close observation of laryngeal function in vivo using his newly created laryngoscope, he presented his findings to the Royal Society of Medicine in his historic paper, “Observations on the Human Voice.” This paper, which was published in the same year (1855), paved the way for the mainstream use of the laryngoscope as a primary diagnostic tool in clinical practice. Although he invented the laryngoscope in an effort to understand the mechanisms underlying the art of singing, he was unaware of the monumental impact that his invention would have in the field of medicine.

In addition to his scholastic achievements, Garcia was a renowned and much sought after teacher of singing. Aside from his two sisters, who both became very famous singers, his notable pupils included Mathilde Marchesi, Charles Battaille, Sir Charles Santley, Julius Stockhausen, Antoinette Sterling, Henrietta Nissen, Johanna Wagner (niece of composer Richard Wagner), Catherine Hayes, Anna Schoen-René, and most famously, Jenny Lind. Many of his students were not only great singers but also went on to become great teachers and scholars. Marchesi, for example, became an influential teacher and passed on Garcia’s teachings to her own students, which included prominent singers such as Emma Eames, Nellie Melba, and her own daughter Blanche Marchesi. Julius Stockhausen, another of Garcia’s students, published a famous treatise on singing based on Garcia’s method called *Gesangsmethode* in 1884. Thus Garcia’s legacy and teachings were successfully passed on by his students.

Despite the Garcia School’s prominence during the early 19th century, it was equally rivaled by the popularity of the Lamperti School. The Lamperti School, which was founded by Francesco Lamperti, was a strong proponent of the old Italian tradition of singing. Unlike Garcia II, whose aim was to develop a scientific basis of good singing by investigating the physiology of the vocal folds during the act of singing, the aim of the Lamperti School was simply to preserve the old Italian vocal tradition that had been passed down to their generation. Instead of giving particular importance to the physiology of the vocal mechanism, the Lamperti School utilized traditional and sometimes ambiguous terminology to evoke the sensations and images of great singing. As the central philosophies and pedagogical approaches of the two schools of singing diverged from one another, there was, naturally, quite a bit of rivalry between the two schools.

Francesco Lamperti had a long and illustrious career as one of the great singing teachers of his time. He taught at the Milan conservatory for a quarter of a century and later taught privately. As a highly sought-after teacher, he attracted students from all around the world and produced a great number of outstanding singers. Among his students were Sophie Cruvelli, Emma Albani, Désirée Artôt, David Bispham, Italo Campanini, Teresa Stolz, Marie van Zandt, Maria Waldmann, and Herbert Witherspoon. In addition to being a renowned teacher, he also wrote several treatises and manuals on the art of singing including his famous *Guida teorico-pratica-elementaire per lo studio del canto* (F. Lamperti, 1864). Despite writing several works detailing his techniques, methods, and exercises for great singing, he did not concern himself with the specifics of vocal anatomy or physiology. Often, sections in his treatises on anatomy and physiology were extracted from other works. However, this did not weaken his teaching techniques in the least. His influence as a teacher in Italy was so great that he was awarded the honor of “Commander of the Crown of Italy” for his services to music.

Francesco Lamperti's son, Giovanni Battista Lamperti, also became an equally renowned teacher of singing and took over the Lamperti School after his father died. He studied with his father in Milan while at the same time serving as a piano accompanist for his father's students. Thus, he became very well acquainted with his father's teaching techniques and style. During his career as a prominent teacher, he taught in various places including Milan, Paris, Dresden, and Berlin. Like his father, he produced numerous first-rate singers. Among his pupils were Irene Abendroth, Marcella Sembrich, Ernestine Schumann-Heink, Paul Bulss, Roberto Stagno, and Franz Nachbaur.

Although he published several works regarding vocal pedagogy, he is primarily remembered for his book, *Vocal Wisdom*, which was transcribed by Lamperti's pupil William Earl Brown (G. B. Lamperti & Brown, 1957). The book is comprised of a collection of maxims that are organized into several loosely tied chapters that each offer a wide array of insights and advice. The primary goal of this book was not to provide exercises and quick fixes to vocal problems, but rather to convey a philosophical disposition through the use of sage proverbs and beautiful analogies. Although the book was compiled by Lamperti's student and was not his direct work, it nevertheless provides valuable insights into the teachings and methodology of the Lamperti School.

Like his father, G. B. Lamperti was not as concerned with the physiological aspects of singing, instead focusing on describing the sensations and timbres associated with good singing. Both Lampertis frequently utilized traditional vocal terminology and phrases that were sometimes quite vague and abstract compared with the specific scientific terminology used by the Garcia school. However, although these terms and phrases may seem enigmatic to those who require precise physiologic or acoustic reference points, they offer profound meaning to those who are already acquainted with the conventions of singing. Despite using ambiguous terminology and phrases with more than one clear interpretation, the Lamperti School enjoyed immense popularity.

As mentioned earlier, due to the differing philosophical approaches of the two schools, there was inevitably some rivalry between them. For instance, G.B. Lamperti, in his book *Vocal Wisdom,* openly declared his aversion to "voice doctors" who, according to him, teach the singer some new trick that often undermines the innate power and control given by nature (G. B. Lamperti & Brown, 1957, p. 21). Although it is not clearly specified in the passage, there is little doubt that he was referring to Garcia. Furthermore, Lamperti asserted that when Jenny Lind lost her voice, she was not able to regain it even after a prolonged study with Garcia and she was only able to regain it when she went home and worked it out herself (G. B. Lamperti & Brown, 1957, p. 21). However, Nathaniel Parker Willis, in his book *Memoranda of the Life of Jenny Lind*, tells the story a bit differently. According to Willis, when Lind first arrived in Paris in order to study with Garcia, her voice was on the brink of extinction, and she was asked to take three months of complete vocal rest before engaging in any further lessons (Willis, 1851, p. 19). In fact, Jenny Lind herself said in a letter home: "My voice has in this short time changed so significantly for the better that it borders on the incredible. I am delighted beyond words with Garcia's care of both me and my voice that I have developed a quite healthy desire for singing" (as cited in Pilotti, 2009).

On the opposite side, some of Garcia's friends publically expressed their animosity toward the Lamperti School. For example, Charles Lunn, a highly regarded voice teacher from Manchester, England, who taught at the Royal Academy of Music and a friend of Manuel Garcia, said in a letter to the editor of *The Music Standard*:

> It is scarcely just for me to draw upon raw experience of boyhood years, but from what I heard of Signor Lamperti's pupils, I certainly thought his method based upon entirely false, and in great degree vicious, notions of voice. It was a deep repugnance felt at the modern Italian school that made me throw my uttermost energy into the scientific corroboration of Garcia's truths. (Lunn, 1862, n.p.)

These examples clearly illustrate the extent of the rivalry between the two schools. However, some scholars argue that while the pedagogical method and focus of the two schools differed greatly, their overarching ideals were not entirely incompatible. One example that reinforces this argument is that Marcella Sembrich, who studied in the Lamperti School, and Anna Schoen Rene, who studied in the Garcia school, remained lifelong friends and colleagues at the Julliard School of Music in New York (Pilotti, 2009, p. 14). The lasting friendship of these two singers from rival schools reveals the fact that the rivalry between the schools may have been a result of the competitive atmosphere brought about by the 19th century golden age of singing.

The final golden age, according to Pleasants, occurred between 1880 and the First World War (Pleasants, 1966). The leading singers of this period included, among others, several of the pupils of both the Garcia and Lamperti Schools such as Nellie Melba, Marcella Sembrich, Emma Eames, and Ernestine Schumann-Heink. Although other scholars have classified an additional golden age lasting throughout the middle of the 20th century, Pleasants (1966) simply describes this period as an “Italian Afterglow” (p. 284). One thing that all the golden ages shared was that they were all dominated by singers of the Italian tradition. Right from the time of Caccini, the Italian vocal tradition remained the preeminent style of singing. This fact is further evidenced in the historical literature, in which other national styles of singing are frequently measured against the superior Italian tradition.

Nevertheless, the years between the individual golden ages of singing were marked by periods of transition and stylistic change. During these transitional times, there were strong sentiments that the art of singing was on the decline. Although this feeling happened in all the transitional periods, it was particularly intense toward the latter half of the 19th century between the third and fourth golden ages of singing. Due to the stylistic changes brought about by composers such as Wagner, many people believed that the art of good singing would be lost forever. Composers, singers, and teachers alike did not hesitate to express this view in their writings. Both Garica II and Francesco Lamperti also criticized this change. They lamented the disappearance of castrati, the stylistic changes in opera favoring vocal declamation over floridity and agility, as well as the simplification of vocal lines in exchange for elaborate orchestral effects (Stark, 1999, pp. 220-221). The term *bel canto* was introduced into the literature during this time to lament the loss of the older style of singing.

As mentioned earlier, the purpose of this document is to illuminate the acoustic principles that are representative of the *bel canto* style of singing. However, this cannot be done without first establishing a general scientific groundwork of the art of singing. Therefore, in the following chapters, the primary aim is to describe the various physiological and acoustic principles underlying the process of singing. Once this groundwork has been laid, it will then be possible to synthesize the teachings of the *bel canto* style with the respective acoustic phenomena.

Chapter 3: THE PROPERTIES OF SOUND

Physical Foundations of Sound

The process of singing has fascinated many individuals throughout the course of history. This fascination and curiosity was further heightened by the fact that the vocal mechanism is hidden away from plain sight. In fact, Garcia's invention of the laryngoscope was fueled by his curiosity to comprehend the complex vibrations of the vocal folds. Although Garcia's invention was a huge step forward to understanding the complex process underlying voice production, the progression of science and technology has allowed for many more discoveries to be made in the field since Garcia's time. Thus, a foundational understanding of modern voice science helps illuminate the acoustic foundations of the *bel canto* technique. However, in order to fully understand the mechanics of voice production as it relates to singing, it is necessary to first understand the basic physical characteristics of sound itself. This chapter aims to build a foundation on the properties of sound so that it can be later applied to voice production.

The definition of sound has frequently been the subject of philosophical debate mainly due to a popular question set forth by George Berkeley in 1710 and paraphrased by numerous later scholars (Berkeley, 1710/1957). The question, with several variations, goes something like this: If a tree falls in an uninhabited island, does it make a sound? Although this question may have originally been intended to probe readers to examine the relationship between perception and reality, in a strictly scientific view there is no controversy within this question. When this question was posed in the April 1884 issue of *Scientific American,* it was clearly stated that sound is a sensation that is recognized at our nerve centers. Thus, although there would be physical vibrations created by the falling tree, it would not be recognized as sound without any ears to perceive it ("Correspondence," 1884, p. 218). For the purposes of this document, the focus is on this scientific viewpoint rather than the philosophical aspect of perception and reality.

In accordance with the scientific perspective, sound can essentially be described as an oscillation of pressure that is propagated through a medium and perceived by our ears. From this definition of sound, it is already clear that a medium of travel and a perceiver are both necessary components for sound. However, upon closer scrutiny of this definition, the perceptive reader may realize that the oscillation of pressure must be initiated by something. In order for a pressure oscillation to occur, it must be generated by a vibrating object that is powered by an energy source. Therefore, in order for any sound to be perceived, there must be at least four individual prerequisite components: an energy source, a vibrating object, a medium, and a perceiver (McKinney, 2005, p. 20). First, the energy source must directly set the vibrating object in motion. Next, this vibrating action of the object displaces the surrounding molecules in the medium and creates an oscillation of pressure. This oscillation of pressure is then propagated through the molecules in the medium until it reaches a perceiver, such as our ears, which subsequently analyzes this physical pressure wave and sends it to our brain, which finally translates it into the sensation of sound.

Sound waves are classified as longitudinal waves; that is to say, the propagation of the oscillation of pressure through the medium occurs in a longitudinal fashion, much like a chain of dominos. Unlike transverse waves (e.g., light waves) in which the molecules being displaced move perpendicular to the direction of the energy being transmitted (e.g., a rope being moved up and down to create a wave moving forward), the molecules in a longitudinal wave move parallel to the direction of energy movement (e.g., dominoes falling forward to propagate a wave in the same direction). However, despite the fact that transverse waves and longitudinal waves have different mechanisms of propagation, they are both graphed in a similar fashion. Generally, both are graphed on a coordinate plane showing amplitude as a function of time. Nevertheless, it must be emphasized that while sound waves may appear as transverse waves when graphed in this way, one must keep in mind that they are still longitudinal waves and propagate in a way that is different from transverse waves.

As mentioned before, sound waves are oscillations of pressure that move through a medium. But how exactly does this happen? In order to understand the mechanics of sound propagation, it is necessary to describe the principles of a medium. Although sound waves can propagate through several different mediums, the most common medium is air. As with all mediums, the air around us is comprised of small particles of matter known as molecules. These molecules all tend to remain a certain distance apart from one another such that, if pulled apart or pushed together, they will always return to their original point of equilibrium. This principle is known as the elastic property of matter (Vennard, 1967, p. 1). Therefore, as the vibrating object moves back and forth in the medium, the adjacent air molecules are compressed forward into the next set of molecules and initiate a rippling effect. It must also be mentioned that whenever air molecules are compressed together, it also implies that they are at the same time being pulled apart, or rarified, from the neighboring air molecules on the other side. Thus, compression and rarefaction are both essential to the propagation of sound waves. The areas where air molecules are densely compressed together results in increased air pressure, while the areas where air molecules are rarified from one another results in decreased air pressure. It is this cycle of compression and rarefaction that is commonly referred to as the oscillation of pressure, and it is in this manner that sound waves are transmitted through the medium (Rosen & Howell, 1991, p. 8).

Musical and Non-musical Sounds

Now that a foundation of sound has been laid out and the mechanism of wave propagation as well as the manner of graphical representation has been established, it is possible to categorize the various types of sound. All sound can be classified into two distinct categories: periodic and aperiodic sounds. Periodic sounds, as their name implies, are comprised of a pattern of pressure oscillations that repeat periodically. Aperiodic sounds, on the other hand, are comprised of random pressure oscillations with no particular repeating patterns (Rosen & Howell, 1991, p. 25-29). Sound waves with periodicity are perceived to have pitch and are consequently labeled as musical tones. However, sound waves without periodicity are, with some exceptions, generally not perceived to have pitch and are therefore labeled as noise. The perception of pitch, along with other characteristics of a musical tone, is discussed in greater detail in the next section concerning the elements of a musical tone. However, in order to fully appreciate the physical and perceptual characteristics of a musical tone, we must first understand the anatomy of a periodic sound wave.

As mentioned before, a periodic sound is synonymous to a musical tone. The simplest of all musical tones is known as a sinusoidal or sine wave. A sine wave is the result of a certain predictable vibratory pattern known as simple harmonic motion or uniform circular motion (Rosen & Howell, 1991, p. 14-19). This motion can best be illustrated by the swinging of a pendulum on a grandfather clock. If the displacement of a moving pendulum were to be graphed as a function of time, the resulting waveform would be sinusoidal in shape. Perceptually, a sine wave is heard as a thin, clear tone. In fact, sine waves are also known as pure tones because they are comprised of only a single frequency. True sine waves are rarely heard in everyday life. Instead, much of the sounds we hear and experience daily are complex sounds comprised of more than one frequency. However, a common object that produces a sine wave is a tuning fork. The tines of a tuning fork, when struck, move in simple harmonic motion just like the swinging of a pendulum. Thus the resulting sound is a pure tone comprised of a single frequency.

Although we now know that simple tones are produced as the result of simple harmonic vibrations and are referred to as pure tones since they are comprised of a single frequency, we still have not clearly outlined the composition or characteristics of a complex sound. Much of what we know about complex sounds was due to the work of the French mathematician, Jean Baptiste Joseph Fourier. Fourier demonstrated that all complex sounds were actually a sum of several simple tones. He developed a sophisticated mathematical procedure that is still used frequently today (although nowadays primarily through the aid of computer programs), by which complex sounds are separated into their simple components. In honor of his contributions, this method is known as Fourier analysis (Rosen & Howell, 1991, p. 117). Thus, simple sinewaves can be considered the fundamental building blocks of all complex sound, including both musical tones as well as noise. While sine waves themselves are periodic in nature and can be considered musical tones, they can be combined together in a random irregular fashion which results in a complex aperiodic sound, also referred to as an inharmonic complex or noise. On the other hand, when they are combined in such a way that the frequencies of the individual sine waves are related to each other mathematically, the resulting sound is a complex periodic wave, also referred to as a harmonic complex or a musical tone. The specific mathematical relationship of the individual pure tones in a harmonic complex is actually what gives the tone its unique quality. However, this concept is elaborated upon and made clearer in the discussion regarding timbre.

Elements of a Musical Tone

There are five main dimensions by which a musical tone can be characterized. Each dimension consists of a physical parameter as well as a perceptual counterpart. Generally, the physical and perceptual aspects are related in such a way that any alterations made to the physical parameter of the wave will directly influence the perceptual aspect of the sound. Additionally, there are a few interrelationships among these dimensions.

The first dimension of a musical tone is the pitch. The perceived pitch of a periodic sound is directly correlated to the rate at which the pattern of oscillations repeat, which is referred to as the frequency of the wave. Simply put, frequency is the number of cycles that occur in the span of one second. By increasing the frequency of a periodic sound, the pitch is also increased. Thus, the physical parameter of frequency is directly related to the perceived pitch of the musical tone. In other words, doubling the frequency results in an upward octave shift in pitch, and halving the frequency results in a downward octave shift in frequency (McKinney, 2005, p. 23). In the case of singing, the frequency of a sung note is determined by the number of cycles the vocal folds vibrate in one second. When a singer sings the note A-440, the number 440 denotes the number of times the vocal folds are oscillating each second. This oscillation of the vocal folds subsequently causes the surrounding air molecules to vibrate at 440 cycles per second, and also causes the eardrum of the listener to be set into vibration at 440 cycles per second. The SI unit (*Système International d'Unités*) used to measure the number of cycles per second, or frequency, is known as hertz (abbreviated Hz), named after the scientist Heinrich Hertz.

Since the pitch of the musical tone is generally dependent on its frequency, it is tempting to assume that only periodic sounds can evoke a sense of pitch. While it is true that periodic tones do evoke a sense of pitch and aperiodic sounds are usually perceived as noise, it must be remembered that pitch is a perceptual sensation which is created in the brain as a response to the physical frequency of the sound. Since the perception of pitch is formulated by our brain, it is possible to trick our brains and elicit the sensation of pitch even with sounds that are technically aperiodic in nature. The purpose of highlighting this subtlety is to emphasize that the sensation of pitch should not be mistaken for the physical parameter of frequency. Thus, while these two terms are sometimes mistakenly used interchangeably, it is important to realize the difference.

A second property of musical tones is their amplitude. The amplitude of a sound describes the magnitude by which the air molecules in the sound wave are being displaced. Generally, amplitude is denoted by the amount of pressure change in the air molecules transmitting the sound wave Rosen & Howell, 1991, p. 24). Going back to the example of the pendulum, the amplitude is analogous to the width of the arc of the pendulum. Although it may at first seem counterintuitive, the width of the arc through which the pendulum swings has no bearing on the frequency of its oscillation. In other words, no matter how high or low the arc of the pendulum, it will take the same amount of time for it to complete one cycle. Thus the law of the pendulum demonstrates that frequency and amplitude are independent of one another.

Another term that is closely related to amplitude is intensity. The intensity of a sound refers specifically to the energy expended per second measured over a particular area (Ferrand, 2007, p. 35). The intensity of a sound wave is directly determined by its amplitude. Both amplitude and intensity are physical parameters; however, the perceptual correlate of these two parameters is loudness. Just like pitch, loudness is a sensation that is created in the brain. However, the relationship between the physical parameters of amplitude and intensity are not linearly related to the perceived loudness. The main reason for this is due to a mechanism of the inner ear called amplitude compression. In short, amplitude compression means that a doubling of the intensity does not result in a doubling of the perceived loudness. This is because the ear provides greater compression for high intensities than for low intensities. Because of this phenomenon, amplitude compression allows the human ear to be able to hear a wide range of intensity levels (Rosen & Howell, 1991, p. 248-249).

Since amplitude compression increases the range of intensities that can be perceived before approaching the threshold of pain, the perception of loudness therefore is not as sensitive to small changes in intensity, particularly in the higher intensities. For example, a sound must be 10 times as intense to be perceived as twice as loud but 100 times as intense to be perceived as three times as loud (Ferrand, 2007, p. 36). The decibel scale (abbreviated dB) was conceived in order to measure sounds in a way that accounts for this non-linear relationship. The decibel scale, named after Alexander Graham Bell, is a logarithmic scale that eliminates the problem of having to use extremely large numbers to express the intensity of a sound. Unlike a linear scale, in which the distance between each increment remains constant, a logarithmic scale functions in such a way that the distance between each increment grows successively larger. In essence, the decibel scale is a base 10 logarithmic scale in which each step represents an increase by a factor of 10. In this way, the decibel scale allows large number of units of intensity to be expressed in a simple, condensed manner.

A third characteristic of musical tones is their duration, which is one of the most straightforward properties of sound. In essence, duration involves the amount of time the sound persists. While it is an easy concept to understand, it is nevertheless one of the most important aspects of music, for without duration, there would be no sense of rhythm or meter. Generally duration is measured in seconds. It must be noted that in some instruments, duration and intensity are directly related. That is to say, a greater initial amplitude will result in a greater duration (Vennard, 1967, p. 4). For example, a tuning fork that is struck with more power will sound for a greater amount of time than one that is struck lightly.

A fourth dimension by which musical tones can be characterized is timbre. Timbre, in a nutshell, refers to the quality of the tone. The timbre of a tone is what allows us to distinguish it from another musical tone of the same pitch. For example, a flute and a trumpet playing the same note can be easily distinguished by their unique timbre. Like pitch and loudness, timbre is also a perceptual quantity that is also influenced by physical parameters. In order to understand the physical parameters that influence the timbre of a musical tone, we must recall the definition of a complex tone. As mentioned earlier, a complex tone is comprised of two or more simple sine waves, each with their own frequency and intensity levels. It is this specific combination of frequencies and intensities of the individual sine waves that contribute to the perceived timbre of the tone. Generally, the frequency of the lowest sine wave in the harmonic complex, which is known as the fundamental frequency or first harmonic, determines the perceived pitch of the tone. The rest of the sine waves, from the second harmonic onwards, are referred to as overtones. In a complex musical tone, the overtones are all multiples of the fundamental frequency. If this mathematical relationship is not present, the resulting sound will be an inharmonic complex and be perceived as noise (Rosen & Howell, 1991, p. 120). Thus, while the pitch of a complex tone is generally determined by the fundamental frequency, the primary cues for timbre are the number, selection, and strengths of the overtones.

Finally, a fifth dimension by which musical tones can be characterized is sonance. While the concept of sonance is not embraced by all scholars, it is championed by a few prominent scholars such as William Vennard, Carl Seashore, and James Stark. Sonance is similar to timbre in that it involves a fusion of sound. Whereas a tone's timbre is the result of a simultaneous fusion of sine waves, the sonance of a tone is determined by a successive fusion of the slight changes in pitch, intensity, and timbre (Stark, 1999, p. 145). The sonance of a tone gives it a more natural quality. A singer's vibrato is a great example of sonance. Vibrato involves slight fluctuations of pitch, intensity, and timbre that fuse together to create a distinct vocal color. Thus, it can be argued that sonance is an important dimension of musical tones.

Now that a fundamental basis of the properties of sound has been established, it is possible to start focusing on the mechanics of voice production. The concepts discussed in this chapter, although somewhat dry and technical, are frequently referenced and, in essence, serve as building blocks in the upcoming chapters on vocal mechanics and aural perception. Finally and most importantly, this information is integral to understanding the acoustical principles implicit in the major tenets of *bel canto*.

Chapter 4: THE PRINCIPLES OF SINGING

In order to illuminate the acoustical principles inherent in the traditional and empirical teachings of *bel canto*, it is first necessary to obtain a working understanding of the anatomy and mechanics of the human voice. Only after first establishing a modern scientific perspective on the aspects of sound production, physiology of singing, and perceptual aesthetics can these aspects be analyzed through the lens of *bel canto*. Thus, this chapter aims to provide the reader with a modern scientific understanding of the mechanics of singing. The concepts described in the previous chapter regarding the properties of sound are frequently employed in this chapter to better illustrate the manifold physiologic processes underlying the art of singing.

The Respiratory Mechanism

Although the concept of breathing is often perceived as a natural and perhaps automatic process, many voice teachers and pedagogues agree that breathing is a core principle of singing that has a significant impact on several other aspects of vocal technique. While it is true that proper breathing technique alone does not necessitate good singing, insufficient breathing technique may directly lead to poor intonation, inconsistent tone quality, inadequate phrasing, unwanted tension, and a multitude of other problems. It is for this reason that breathing plays such a significant role in singing.

Breathing for speech and breathing for singing, while slightly different, share similar underlying physiological processes. Furthermore, since breathing for singing essentially builds upon the processes of speech breathing, it is beneficial to first consider the basic mechanics of speech breathing and subsequently apply it to breathing for singing. However, before focusing on the mechanics of speech breathing, it is important to build a foundation of the anatomy of the respiratory mechanism.

The respiratory system is comprised of three main units: the upper respiratory system, the lower respiratory system, and the chest wall system (Ferrand, 2007, p. 69). The oral, nasal, and pharyngeal cavities make up the upper respiratory system. Some classifications also include the larynx as part of the upper respiratory system. The lower respiratory system is comprised of the trachea, the bronchial system, and the lungs. The chest wall system is comprised of the thoracic cavity, which is bounded by the vertebrae, ribs, sternum, diaphragm, and pectoral girdle, as well as the abdominal cavity, which is bounded by the vertebrae, diaphragm, and pelvic girdle. Thus, the diaphragm separates the two cavities of the chest wall system by serving as the floor of the thoracic cavity and the roof of the abdominal cavity. The chest wall system essentially houses a variety of organs, including the lungs. However, the importance of the chest wall system in respiration is due to the manner in which it functions together with the lungs.

The lungs, which are essentially air-filled elastic sacs, are coated by a thin lubricated tissue called the pleura. The same tissue also coats the inside of the thoracic cavity. The fluid-filled space in between these two pleura is known as the pleural cavity. The pressure inside the pleural cavity is usually negative, meaning that it is much less than the atmospheric pressure outside the body (Ferrand, 2007, pp. 78-79). This is particularly significant because the negative pressure plays a pivotal role in attaching the lungs to the thoracic cavity. The pleural cavity attaches the lungs to the thoracic cavity in much the same way as suction cups attach to a smooth surface. When pushed on to a smooth surface, the suction cup expels much of the air inside of it, creating a partial vacuum. In other words, the air pressure inside the suction cup is much lower than the atmospheric pressure outside it. Consequently, the outside air pressure, which is greater, pushes the suction cup on to the surface, causing it to stick. In the same manner, since the pleural pressure—the pressure in between the lungs and the thoracic cavity—is negative, the lungs are stuck to the chest wall just as the suction cup sticks to a smooth surface. Thus, any movement of the chest wall directly causes movement of the lungs. Figure 1 shows the basic structure of the respiratory system, revealing the manner in which the diaphragm is attached to the lungs.

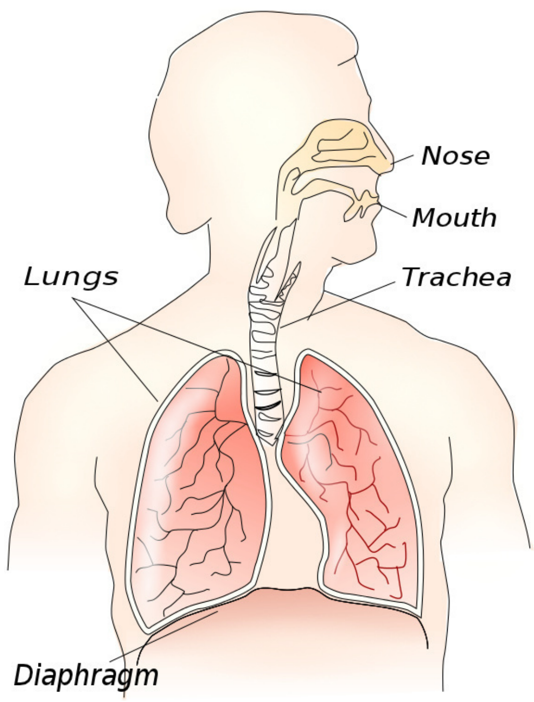

Figure 1. Basic Structure of the Respiratory System (n.d.)
Note: In the public domain.

It is important to understand that, although the pleural cavity connects the lungs and chest wall together, both the lungs and chest wall are held together in a state of elastic tension. In other words, if the adhesive force of the pleural cavity were disengaged, the lungs, in isolation, would have a tendency to collapse inward while the chest wall, in isolation, would have a tendency to expand outward. Thus, it is the negative pressure of the pleural cavity that holds these two opposing forces together in a state of elastic balance. Furthermore, this pleural linkage between the lungs and thorax plays a crucial role during the mechanics of breathing as it causes the lungs and chest wall to move synergistically as one unit. This concept is further elaborated upon in the discussion about the physiology of respiration.

In order for air to reach the lungs, it must first enter through the nose or mouth and travel down through the larynx and trachea. The trachea, or windpipe, carries the air down into the bronchial system, sometimes also referred to as the bronchial tree. It is referred to as such due to its numerous subdivisions, which resemble the branches and twigs of a tree. Beginning with the first subdivision of the trachea into two mainstream bronchi that each enter one lung, there are approximately 28 orders of subdivision in the bronchial tree that finally conclude with the tiny respiratory bronchioles (Seikel, King, & Drumright, 1997, p. 63). At the end of these respiratory bronchioles are thin-walled, air-filled, microscopic structures known as alveoli (singular: alveolus). The number of alveoli in the adult lung is estimated to be somewhere between 300 to 750 million (Ferrand, 2007, p. 71). These alveoli and the capillaries that surround them are the primary facilitators of the oxygen and carbon dioxide gas exchange. In other words, it is at this location that oxygen enters into the blood stream and carbon dioxide exits out of the blood stream. Thus the alveoli play a very significant role during the respiratory process.

Despite the significance of the alveoli in the lungs, they would not be able to fulfill their role without the active expansion and contraction of the lungs. However, since the lungs are spongy structures that cannot expand or contract by themselves, their movement is regulated by the various muscles surrounding them. Although there are numerous muscles that can potentially be involved in respiration, there are a few that play critical roles in respiration. Perhaps the most important muscle of respiration is the diaphragm. As mentioned earlier, the diaphragm separates the thorax from the abdomen. It is a thin dome-shaped muscle that acts as the floor of the thoracic cavity as well as the roof of the abdominal cavity. The diaphragm and the lungs are not directly connected; instead they are held together by the negative pleural pressure in between them. This is the concept that was mentioned earlier—any movement of the thoracic cavity results in movement of the lungs as well. Thus, when the diaphragm muscle contracts, its dome shape flattens downward and consequently expands the lungs down vertically with it. When the diaphragm relaxes, on the other hand, the lungs revert back to their original size and shape.

Another significant set of muscles that aid in respiration are the external and internal intercostal muscles. These muscles run in between the ribs and assist in expanding and contracting the ribcage. The external intercostals work together with the diaphragm and increase the size of the thoracic cavity by expanding the ribcage upward and outward during inhalation (Ferrand, 2007, p. 75). On the other hand, the internal intercostals work to decrease the volume of the thoracic cavity by pulling the ribcage downward during exhalation. Finally, the last crucial set of muscles involved in respiration are the abdominal muscles. , which are comprised of four individual muscles that function together as a unit: the external oblique, internal oblique, rectus abdominis, and the transverse abdominis. These muscles all share the same purpose of compressing the abdomen, thereby pushing the diaphragm further upwards during exhalation. This causes the thoracic cavity to decrease in volume. Although there are several other accessory muscles in the neck, thorax, and abdomen that can be involved in respiration to varying degrees, the key muscles are the diaphragm, intercostals, and the abdominals.

Now that the basic structure and components of the respiratory system have been laid out, the manner in which these components function together can be discussed. The process of respiration, in essence, is facilitated by manipulating the various pressures in the body. In order to gain a deeper understanding of the mechanics of respiration, it is essential to appreciate the relationships between airflow, pressure, and volume. According to Boyle's Law, pressure and volume are inversely proportional (Ferrand, 2007, pp. 11-12; Seikel et al., 1997, pp. 36-37). In other words, when the volume of an enclosed container is decreased, the pressure inside it increases. This increase in pressure can be explained by the elastic property of matter, described in the previous chapter, which dictates that molecules generally tend to remain a certain distance apart from one another. Thus, when the volume of the container is decreased, the molecules are compressed and have less space to move around, which in turn causes more frequent collisions among themselves and the walls of the container. On the other hand, when the volume of the container increases, the pressure inside decreases since the molecules have more space to move about and collide less frequently. Finally, it must also be pointed out that the flow of air always occurs from regions of high pressure to regions of low pressure. This phenomenon occurs due to the tendency of gasses to find an equilibrium within the two pressure gradients (Ferrand, 2007, p. 11). In other words, the air molecules will move from places of higher pressure to places of lower pressure in order to spread themselves out equally.

Equipped with this basic understanding of the relationships between airflow, pressure, and volume, it is now possible to describe the process by which the body regulates the act of breathing. While at rest, as previously stated, the lungs and the chest wall are held together in a state of elastic balance. In this rest position, the air pressure inside the lungs, also referred to as the alveolar pressure, and the atmospheric air pressure outside the body are more or less equal. However, since the pressures are equal, the air cannot flow in or out of the body. In order to draw air into the lungs, the body must first create a pressure difference by decreasing the alveolar pressure inside the lungs. If the air pressure inside the lungs is lower than the pressure outside the body, the air should naturally flow into the body, as air always flows from regions of high pressure to regions of low pressure. The body makes this pressure decrease happen mainly by contracting the diaphragm and the external intercostal muscles. When these muscles contract, the volume of the thoracic cavity, and consequently the volume of the lungs, increases. As the volume of the lungs increases, Boyle's law dictates that the pressure decreases. Thus, the newly created pressure difference causes the outside air to flow in through the nose or mouth and enter the lungs. In summary, air is drawn into the body when the muscles of inspiration (diaphragm and external intercostals) contract and expand the lungs. The increase in lung volume causes a decrease in alveolar pressure, which in turn causes the outside air, which has greater pressure, to flow into the lungs, which has less pressure.

Up till this point, only the mechanics of inspiration have been discussed. However, respiration consists of both inspiration and expiration. As soon as the air enters the lungs, the inside alveolar pressure and outside atmospheric pressure begin to equalize. Once they reach equilibrium and the pressure difference drawing the air into the lungs disappears, the airflow temporarily stops for an instant. At this point, it must be emphasized that since the lungs and thoracic cavity are being held in an expanded state by the external intercostals and the diaphragm, they are resisting a strong elastic recoil force to return to their original state, much like a stretched rubber band. When these muscles relax, the thoracic cavity and lungs recoil back into their resting state. This recoil action causes the air inside the lungs, which by this time has been exchanged from oxygen to carbon dioxide inside the alveoli, to be pushed outside. Once the thoracic cavity has returned to its resting state and the carbon dioxide has been expelled from the lungs, the respiratory cycle begins again. Thus, in quiet breathing (i.e., vegetative breathing), the cycle of respiration is comprised of an active inspiratory process, in which the diaphragm and external intercostals contract and expand the thoracic cavity, and a passive expiratory process, in which the relaxation of the diaphragm and external intercostals cause the thoracic cavity to recoil back to its original size. From birth, this cycle of inhalation and exhalation is controlled subconsciously by our brain via the central nervous system (Ferrand, 2007, p. 80).

While it is true that the respiratory cycle mentioned above is sufficient for vegetative breathing, this type of breathing is insufficient to support the demands of the speaking or singing voice. One of the central differences between vegetative breathing and breathing for speech or singing is the way in which air intake is determined. For vegetative breathing, the rate and levels of air intake are determined reflexively by the carbon dioxide levels in our blood. For example, when the body is engaged in rigorous physical activity and more oxygen is needed, the respiratory center in the brain sends signals to the inspiratory muscles to adjust the rate and depth of inspiration accordingly. Although this process is usually controlled subconsciously, when speech and singing considerations, such as phrase length and dynamics, are integrated into this process, the act of breathing changes from a subconscious reflexive action into a conscious voluntary action. Simply put, vegetative breathing is a subconscious phenomenon while breathing for speech and singing is a conscious phenomenon.

In addition to the conscious voluntary control, breathing for speech or singing introduces a few notable changes that occur in the mechanics of the previously discussed vegetative breathing cycle. The most overt difference is the shift in the location of air intake. In vegetative breathing, the air is generally inhaled through the nostrils in order to moisten the incoming air as well as to trap and prevent dust particles from entering the lungs. However, in breathing for speech or singing, the air is generally inhaled through the mouth in order to increase efficiency of inspiration by shortening the distance to the lungs.

A second major difference between the two types of breathing lies in the amount of air inhaled each cycle. For vegetative breathing, we generally only inhale up to 500 cubic centimeters, which is about 10% of our vital capacity. The vital capacity of our lungs refers to the maximum amount of air that can be exhaled following a maximum inhalation. So, in other words, vegetative breathing utilizes only 10% of the maximum amount of air our lungs can use. In contrast, when breathing for speech or singing, the amount of air inhaled increases depending on the length of the phrase and intensity of the vocalization (Ferrand, 2007, pp. 86, 90).

A third factor that distinguishes the two types of breathing involves the ratio of time taken for inspiration and expiration. In vegetative breathing, the inspiratory phase takes about 40% of the respiratory cycle and the expiratory phase takes about 60% of the respiratory cycle. Thus the ratio of time taken for inspiration and expiration is somewhat balanced. However, while breathing for speech or singing, the time taken for expiration is greatly lengthened compared to inspiration. For speech in particular, the inspiratory phase occupies only 10% of the total respiratory cycle while the expiratory phase occupies the remaining 90% (Ferrand, 2007, p. 89; Seikel et al., 1997, p. 158). The primary reason for this change is that speech and singing occur during the expiratory phase; therefore the exhalation must be controlled in such a way that it spans the length of the phrase being spoken or sung.

This leads us to the final, and most significant difference between speech breathing and vegetative breathing: the involvement of the expiratory muscles. Vegetative breathing, as mentioned earlier, has a passive expiratory phase; in other words, no muscle activity is needed for expiration since the recoil forces revert the lungs back to their original state. On the other hand, expiration while singing or speaking is an active process requiring the action of the expiratory muscles in order to compress the thoracic cavity and lungs. Although the anatomy and physiology of the main expiratory muscles, which consist of the internal intercostal muscles and the abdominal muscles, were mentioned earlier, the reason for their activity in breathing for speech and singing has not yet been addressed. The key reason for an active expiratory phase stems from the need to maintain a relatively constant alveolar pressure in order to sustain phonation. For normal conversational speech, the body must maintain an alveolar pressure of around 5-

10 cm H_2O (Ferrand, 2007, p. 92). Without this subglottal pressure, the constant flow of air that is needed to drive the vocal folds could not be generated. Thus, the muscles of inspiration and expiration are employed to maintain this constant alveolar pressure throughout the duration of the spoken or sung phrase.

Although the breathing cycle for speech or singing has been distinguished from the vegetative breathing cycle by the variances outlined earlier, in order to provide a more comprehensive understanding of this cycle, the major events of this cycle must be summarized. Until inspiration, the mechanics of vegetative and speech breathing are similar; the diaphragm and the external intercostals contract in order to expand the thoracic cavity, which results in a lower alveolar pressure, thus forcing air into the lungs. After this point, however, the underlying mechanics of these two cycles begin to diverge. In vegetative breathing, the diaphragm and external intercostals relax and allow the thoracic cavity and lungs to recoil back into their original state. However, this passive recoil does not generate the constant alveolar pressure that is needed to sustain phonation. Therefore, in speech breathing, after inspiration is complete, the inspiratory muscles do not relax. Instead, they stay contracted in order control the rate that the lungs and thoracic cavity return to their original state, thereby maintaining a relatively constant subglottal pressure.

Although in vegetative breathing, the expiratory phase would be complete once the lungs and thoracic cavity return to their original state, this is not the case for speech breathing. Once the lungs return to their original state and the alveolar pressure inside is equal to the atmospheric pressure outside, the muscles of expiration contract and further compress the thoracic cavity and lungs past their resting state in order to maintain the constant alveolar pressure required to sustain vocal fold phonation. As soon as the spoken or sung phrase is complete, the expiratory muscles relax, causing the overly compressed lungs and thoracic cavity to recoil and expand back to their resting state. In summary, alveolar pressure is held constant throughout exhalation first by the inspiratory muscles, which control the speed at which the lungs and thoracic cavity recoil back into rest position, and subsequently by the expiratory muscles, which further compress the lungs and thoracic cavity until the end of the utterance or sung phrase. This act of delaying the collapse of the thoracic cavity and subsequently compressing it in order to maintain a constant alveolar pressure is known as breath control (Miller, 1986, p. 278). Thus, both the inspiratory and expiratory muscles are essential to maintaining a constant subglottal pressure throughout expiration for speech or singing.

Throughout this discussion, it was made evident that the mechanics of breathing for speech and singing were distinctly set apart from the mechanics of breathing for vegetative purposes. However, the specific differences between breathing for speech and breathing for singing were not fully illuminated. Although research has shown that the respiratory patterns utilized for certain singing styles are indeed very closely related to the respiratory patterns utilized for speech, it has also been demonstrated that there are some differences in the breathing patterns of classically trained singers that set them apart (Hoit, Jenks, Watson, & Cleveland, 1996). These differences primarily stem from the fact that classically trained singers generally utilize greater volumes of air and higher alveolar pressures than the speaking voice in order to achieve the wide range of dynamics and colors demanded by their repertoire. In fact, it has been observed that in some instances classical singers can utilize up to 100% vital capacity of the lungs, maximum ribcage capacity, and maximum abdomen capacity (Watson & Hixon, 1985). In contrast, research by Hoit and colleagues (1996) found that country singers utilize only between 16-35% vital capacity, which is only slightly higher than the range, 14-23% vital capacity, they utilize for conversational speech.

Aside from the sheer volume of air used, the other key difference mentioned by Hoit and colleagues (1996) that distinguishes the breathing patterns of classically trained singers occurs during the transition between inspiration and expiration and the transition between expiration and inspiration. While these transitions are unremarkable in untrained singers, classically trained singers exhibit unique behaviors that optimize efficiency of breath. In classical singers, the period between inspiration and expiration is marked by swift isovolume adjustments of the ribcage and abdomen. In other words, once inhalation is complete, the volume of the abdomen decreases in equal proportion to the increase in ribcage expansion without changing the overall volume of the lungs (i.e., the abdomen is tucked in as the ribcage is elevated). The purpose of this action is two-fold. First, the expansion and elevation of the ribcage places it at a mechanical advantage for making quick expiratory pressure adjustments. Second, the tucking in of the abdomen places the diaphragm at a mechanical advantage for making quick inspiratory pressure adjustments. On the other hand, the period between expiration and inspiration in classically trained singers is marked by a rapid decrease in lung volume immediately preceding inspiration. In essence, there is a quick expulsion of the remaining air before the subsequent inspiration is initiated (Hoit et al., 1996). Thus, these respiratory patterns exhibited by classically trained singers, although requiring greater muscular action, makes it possible for them to efficiently meet the demands of the repertoire.

The Vocal Mechanism

While discussing the physics of sound in the previous chapter, it was stated that sound had four prerequisites: an energy source, a vibrating object, a medium, and a perceiver. Since vocalization is a type of sound, these components can also be applied to singing. The previous section helped to establish a solid foundation of respiratory mechanism, which serves as the energy source in the art of singing. This section builds upon that foundation and explores the precise manner in which the breath activates and powers the vocal mechanism, which initiates vibration in singing. Furthermore, the aspects of registration and resonance will also be addressed through a modern scientific lens. The best way to go about this discussion is to first introduce the various anatomical structures associated with voice production. Although the anatomical discussion includes the structures that are most integral to the art of singing, it is not by any means medically exhaustive, as that is beyond the scope of this document. For a more comprehensive anatomical survey with detailed illustrations, readers are encouraged to refer to Robert Sataloff's (2005) book, *Voice Science*. After this anatomical groundwork has been laid, it will be much easier to explore the complex operation of the vocal mechanism as it relates to singing.

Vocal sound is produced in the vocal organ situated in the neck known as the larynx. Although the larynx is a rather sophisticated organ comprised of several different structures, there are essentially two basic types of structures that are vital to sound production: cartilages and muscles. The cartilaginous structures of the larynx comprise much of its shape, and the muscles connect these cartilages together. The bottom-most layer of the larynx is shaped like a signet ring and is known as the cricoid cartilage. The cricoid sits directly on top of the trachea. The next layer above the cricoid is the thyroid cartilage. The thyroid, unlike the cricoid, does not form a complete ring; instead the shape more closely resembles a curved mask with a notch-like protrusion in the front. This notch, while present in both men and women, is more pronounced in the adult male larynx and is commonly referred to as the Adam's apple. The top of the thyroid cartilage is attached to the epiglottis, a thin leaf-shaped cartilage that covers the airway when swallowing in order to prevent aspiration. Finally, the two pyramid-shaped arytenoid cartilages are located behind the thyroid cartilage and situated on top of the cricoid cartilage. Figure 2 provides a clear posterior view of the larynx and its cartilages. Although there are more cartilaginous structures in the larynx, the aforementioned cartilages are the most vital to voice production.

The muscles of the larynx can be divided into two groups based on their location: intrinsic and extrinsic muscles. The extrinsic muscles, as their name suggests, connect the larynx to structures outside the larynx and are therefore important in controlling the elevation of the larynx. While there are several extrinsic muscles, they can be divided

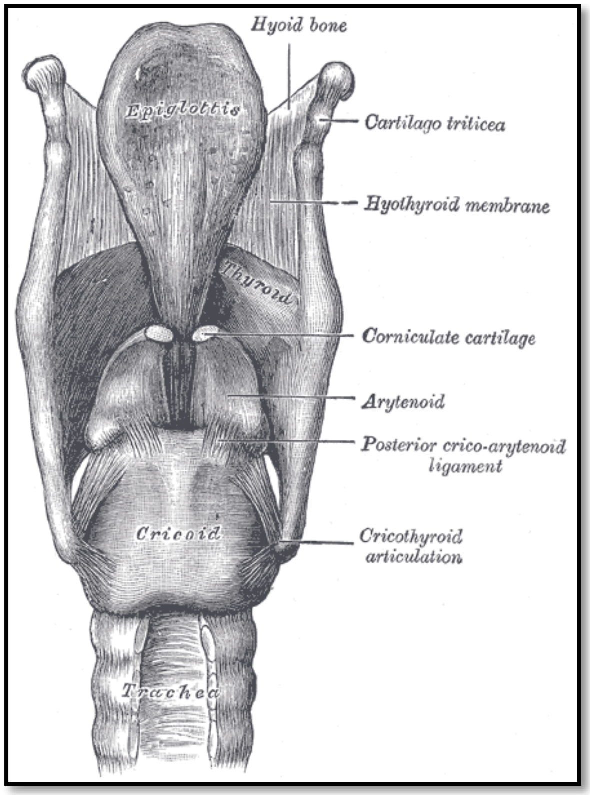

Figure 2. Structure of the Larynx.
Source: *Anatomy of the Human Body* (Gray, Lewis, & Gray, 1918).

into two sections: the suprahyoid muscles which are above the hyoid bone and the infrahyoid muscles which are below the hyoid bone. In general, the infrahyoid muscles, which consist of the thyrohyoid, sternothyroid, sternohyoid, and omohyoid muscles, work together as a group to depress the larynx. In contrast, the suprahyoid muscles, which consist of the digastric, mylohyoid, geniohyoid, and stylohyoid muscles, work as a group to elevate the larynx (Sataloff, 2005, p. 75). Thus, these two groups of muscles are responsible for maintaining a consistent laryngeal position while singing. Although the extrinsic muscles of the larynx have many other useful functions, it is the intrinsic muscles that are absolutely integral to sound production in the larynx.

The intrinsic muscles are responsible for controlling vocal fold adduction (closing), vocal fold abduction (opening), and vocal fold tension in the larynx. Thus, it is these groups of muscles that are responsible for the vibratory patterns of the vocal folds. The nomenclature of the intrinsic muscles of the larynx is very straightforward. Each muscle is named for which two cartilages it connects. The thryoarytenoid (TA) muscle connects the thyroid cartilage to the arytenoid cartilages. This is one of the most significant muscles as it comprises the body of the vocal folds. Thus, the contraction of the vocalis muscle, which is a part of the TA muscle, causes the vocal folds to become shorter and thicker, resulting in a richer and more chest-dominant sound. Relaxation of the TA muscle results in the sound becoming lighter and head-dominant. The specific mechanism underlying this shift in vocal registers will be expounded upon in the subsequent physiological discussion.

Another extremely important intrinsic muscle is the cricothyroid (CT) muscle, which connects the cricoid cartilage to the thyroid cartilage situated on top. Thus, when the CT muscle contracts, it causes the thyroid cartilage to tilt in a rocking motion on the cricoid cartilage. Since the vocal folds, which are a part of the TA muscle, are attached to the thyroid cartilage, this rocking motion of the thyroid cartilage causes the vocal folds to stretch and lengthen. This action is commonly referred to as longitudinal tension. By increasing longitudinal tension of the vocal folds, the pitch of vocalization is also increased. Thus, the primary contribution of the CT muscle in singing is regulating the pitch (Sataloff, 2005, p. 70).

There are two muscles that are responsible for the adduction of the vocal folds. First, the lateral cricoarytenoid (LCA) muscle, which connects the cricoid cartilage to the arytenoid cartilages laterally, regulates the initial adduction of the vocal folds. The contraction of the LCA muscle causes the arytenoid cartilages to rotate away from each other, which results in the closure of the anterior section of the vocal folds. However, although the LCA adducts the anterior glottis, the posterior glottis remains slightly open due to the outward rotation of the arytenoids. In order to achieve a complete closure, the interarytenoid (IA) muscle is subsequently activated. As its name suggests, the IA muscle connects the two arytenoid cartilages together. The contraction of the IA muscle, which is comprised of both oblique and transverse fibers, pulls the two arytenoid cartilages together and closes the posterior gap. This action of the IA muscle, which results in a complete closure of the glottis, is commonly referred to as medial compression (Sataloff, 2005, p. 70). Thus, the LCA muscle plays a significant role in the initial adduction of the vocal folds, while the IA muscle subsequently provides medial compression to fully close the glottis.

Finally, the posterior cricoarytenoid (PCA) muscles are the only muscles that abduct the vocal folds. The PCA, like the LCA, also attach the cricoid cartilage to the arytenoid cartilage. However, unlike the LCA, which connects the two cartilages laterally, the PCA connects them posteriorly. Abduction of the vocal folds is achieved when the PCA contracts and rotates the arytenoid cartilages laterally, pulling the vocal folds away from the midline. The PCA muscles are not only significant to singing and vocalization, but also to breathing. Inspiration can only occur when the PCA muscles contract and abduct the vocal folds. Furthermore, since the PCA is the only vocal fold abductor, a bilateral paralysis of the PCA muscle can lead to a permanent adduction of the vocal folds, resulting in suffocation. Because the integrity of the PCA muscle is extremely important to maintain safe respiration, it is often referred to as the safety muscle of the larynx (Kulkarni, 2012, p. 444).

Now that the basic anatomy of the larynx has been established, it is possible to take a closer look at the underlying physiology of voice production. The voice is first initiated by the vibration of the vocal folds. However, in order for the vocal folds to vibrate, an energy source is required to set them in motion. As mentioned earlier, this energy source comes in the form of breath. In order to inhale, the PCA muscle activates and opens the airway. After the inhalation is complete, the vocal folds are brought together by the LCA and the IA muscles. Once the vocal folds have closed, the subglottal pressure (i.e., the air pressure below the vocal folds) begins to build up until it overpowers the muscular force keeping the vocal folds together and blows them apart. In the past, there were two major theories of phonation: the myoelastic theory and the aerodynamic theory. Up until this point, these two theories were in agreement. However, the way in which the vocal folds return to the closed position and sustain vibration is where the two theories offer divergent explanations. According to the myoelastic theory, it is the adductive muscular tension that brings the vocal folds back together, causing the cycle to repeat continuously. On the other hand, the aerodynamic theory maintains that the vocal folds are brought together by a phenomenon known as the Bernoulli effect. This principle essentially states that increased airflow causes a decrease in pressure. Thus, when the airflow in between the vocal folds increases, the pressure decreases, thereby causing the vocal folds to be suctioned back together and the cycle to repeat. Although these two theories offered different explanations of the mechanics underlying vocal fold vibration, they were not mutually exclusive. In recent years, scientists have proposed that the two theories actually work together to sustain vocal fold vibration and have combined them into the myoelastic-aerodynamic theory of phonation (Ferrand, 2007, p. 130).

The sound produced at the glottal level by the vibrations of the vocal folds are influenced by several factors including the amount of subglottal pressure, the amount of pressure above the glottis, the amount of muscular adduction, the amount of airflow through the glottis, as well as the mass, length, and tension of the vocal folds (Sataloff, 2005, p. 81). For example, the pitch of a sung tone is determined by the frequency of vocal fold vibration. As mentioned in the previous chapter, increasing the frequency of a periodic sound (such as the sound produced by the vibrations of the vocal folds) also increases the perceived pitch. The frequency of vibration of any vibrating object, the vocal folds being no exception, is dependent on its mass, length, and tension. This is the reason why men, who have comparatively longer and more massive vocal folds, have deeper voices than women and children. Although the length and mass of the vocal folds have an influence on the pitch, the primary mechanism by which the singer regulates pitch is through adjusting the tension of the vocal fold cover (Ferrand, 2007, p. 132). Terminology, such as the cover of the vocal folds, will be more easily understood once the microscopic structure of the vocal folds has been described in more detail.

For several years, the histological structure of the vocal folds was relatively unknown. In fact, it was only in 1975 that Minoru Hirano found that the vocal folds have a layered structure (Sataloff, 2013, p. 3). This finding was quite important in explaining the complex vibratory patterns of the vocal folds. Hirano and Bless (1993) proposed that the vocal folds are a multi-layered structure, comprised of five different layers. To make things simpler, they grouped these five distinct layers into a simpler three-layer model. In this model, the outermost layer of the folds, known as the cover, is primarily made up of mucosa (mucous membrane) and soft gelatin-like tissue. The second layer, which they referred to as the transition, is comprised of ligamentous tissue that provides structural support for the folds. The innermost layer of the folds, which they labeled the body, is comprised of the TA muscle. Thus, the three sections of the vocal folds include the cover, transition, and the body.

The discovery of the layered structure of the vocal folds allowed a deeper comprehension of the complex manner in which the folds oscillate during phonation. The slight differences in the mass and texture between each layer of the vocal folds causes them to have different mechanical properties. Due to these contrasting mechanical properties, the passing airflow causes the vocal folds to open and close in a complex wave-like pattern, rather than simultaneously as a whole. In other words, there is a small time lag between the opening and closing of the vocal folds from bottom to top, called the vertical phase difference, and from back to front, known as the longitudinal phase difference (Ferrand, 2007, p. 131). The smooth wave-like motion of the vocal folds, often referred to as the mucosal wave, is a result of these phase differences.

The layered structure of the vocal folds also assists in providing a deeper understanding of the concept of vocal registers. Vocal registration is a topic of heated debate in the field of singing. In fact, both vocal pedagogues and medical experts concede that there is much semantic confusion and controversy regarding it (McKinney, 2005, p. 93; Sataloff, 2005, p. 81). Despite the controversial nature of vocal registers, modern scientific findings have contributed greatly to understanding them. In essence, a register is a series of tones of the same quality produced in the same vibratory manner (Ferrand, 2007, p. 139; McKinney, 2005, p. 93). With this definition in mind, the known registers from low to high include vocal fry, modal voice (sometimes called chest register), falsetto, and whistle. Although the vocal fry and whistle registers have some unique uses in singing, the modal and falsetto registers (and the mix between them) are arguably the two most important to *bel canto*. Over the course of time, scientists and vocal pedagogues have referred to these two registers by various names. Voice pedagogues, such as William Vennard (1967) referred to them as the heavy and light mechanisms (p. 63), while many speech scientists have labeled them as the modal and loft registers. However, in order to avoid further confusion, this document will hereafter refer to them as the modal and falsetto registers.

The underlying mechanism of these two main registers was famously investigated by Janwillem van den Berg (Stark, 1999, p. 82). His findings indicated that the TA muscle was active in the modal register, whereas it was passive in the falsetto register. When the TA muscle, which comprises the body of the vocal folds, actively contracts during vibration, it causes the folds to add more resistance to the longitudinal tension of the CT muscles, thereby causing all the layers to be actively involved in the vibration. This complex pattern of vibration, with the phase differences, is the pattern described earlier. When the modal voice transitions into the falsetto, the TA muscle relaxes and allows the CT muscle to stretch the TA muscle further. Due to the passivity of the TA muscle, only the outer mucosal and ligamentous layers of the vocal folds enter into the vibration. Vennard (1967) describes this process concisely and effectively:

> With the vocalis muscle relaxed it is possible for the cricothyroids to place great longitudinal tension upon the vocal ligaments. The tension can be increased in order to raise the pitch even after the maximum length of the cords has been reached. This makes the folds thin so that there is negligible vertical phase

> difference, no such thing as the glottis opening at the bottom first and then at the top. The vocalis muscles fall to the sides of the larynx and the vibration takes place almost entirely in the ligaments. (p. 67)

So in essence, the primary difference between modal voice and falsetto is the respective activity and passivity of the vocalis muscles. These findings were corroborated by several others soon afterward.

One of the main reasons registration has become such a controversial topic is due to the presence of acoustic registers, also known as secondary registers. Although it was mentioned that one of the principal components that distinguish the primary vocal registers is the change in the distinct pattern of vocal fold vibration, it is not the case for acoustic registers. Shifting between acoustic registers, unlike the primary registers, does not elicit a change in the vibratory pattern of the vocal folds; instead, the perceived shift in quality is due to a shift in the resonances of the vocal tract. Because these resonance shifts are not true registers in the narrowest sense of the term, they are not accepted by all voice scientists. Furthermore, in order to distinguish them from the true laryngeal registers, they are often referred to as resonance registers (Stark, 1999, p. 88). In order to understand how these resonance registers affect the quality of the sound, it is first necessary to understand the relationship between the sound produced by the vocal folds and the shape of the vocal tract.

The Source-Filter Theory

The sound produced at the level of the glottis is very different from the sound that ultimately escapes our lips when singing or speaking. The sound produced by the vibration of the vocal folds resembles an unintelligible buzzing noise, much like the buzzing of a mouth piece of a brass instrument. It receives its unique quality only when it passes through the vocal tract, which consists of the throat, mouth, and nose. The vocal tract, in essence, filters the sound produced at the glottal source by strengthening different areas of resonance, much like the bell of a brass instrument. This interaction between the vibratory source and the shape of the vocal filter is often referred to as the source-filter theory (Ferrand, 2007, p. 199).

The main elements of the source-filter theory are the source, which is the sound produced by vocal folds, and the filter, which is the resonances of the vocal tract. At first approximation, the source spectrum produced by the glottal pulses of the vocal folds resembles a complex tone known as a sawtooth wave, named for its waveform, which is shaped like the teeth of a saw. As mentioned in the previous chapter, all complex tones are a combination of several simple sine waves. The lowest of these simple tones, called the fundamental frequency, determines the perceived pitch of the complex tone. The rest of the tones above the fundamental frequency, called harmonics, are all multiples of the fundamental frequency and determine the timbre of the sound. Thus, while the fundamental frequency is determined by the rate of vocal fold vibration, the buzzing quality of the sound produced by the vocal folds is primarily influenced by its harmonics.

The filter component of the source-filter model is comprised of everything from the space above the vocal folds till the end of the lips (i.e., the vocal tract). The vocal filter, which essentially acts as a tube closed at one end, has its own natural resonant frequencies apart from the harmonics produced by the laryngeal source. As with any tube, the vocal tract's resonances are primarily determined by its size and shape. This occurs because the unique size and shape of the tube influences the way in which sound is reflected as it passes through. In the case of a tube closed at one end and open at the other, these reflections cause a phenomenon known as a standing wave to occur at certain frequencies. Standing waves happen when the sound waves are reflected back in such a way that it causes the reflected wave to interfere perfectly with the incident wave (Hodges & Sebald, 2011, p. 93). These interferences, which happen only at certain frequency areas, cause those frequency areas to be particularly emphasized. Thus, when the sawtooth wave, produced at the level of the glottis, passes through the tube-like vocal tract, the resultant sound wave will have peaks in multiple frequency areas according to the physical characteristics of the vocal tract. These spectral peaks are often referred to as formants. Since the location of these formants are primarily determined by the size and shape of the vocal tract, any slight changes in vocal tract size or shape also shifts the location of the formants. Thus, formant location plays a significant role in our perception of vowels and consonants. This topic will be explored in greater detail in the discussion on articulation.

The original source-filter model assumes that the source frequency, produced by the vocal folds, and the filter formants, generated by the vocal tract, function independently of one another. In other words, the output can only be a linear combination of the individual inputs. Therefore, according to this linear source-filter model, the filter cannot interact with the source to produce any new frequencies. However, recently it has been argued that this assumption is generally not valid since it does not hold true in all situations (Titze, 2008). Although the linear source-filter model has been the primary model of speech analysis in the past, Titze (2008) highlights that it has been recognized all the while that the linear model applies more to men than to children and women. He explains that as long as the dominant source frequencies lie below the formant frequencies of the vocal tract, the source is only mildly influenced by the vocal tract. However, when the lower harmonics of the source cross the formants, more intense nonlinear interactions between the source and filter occur.

The primary mechanism controlling the degree of interaction, according to Titze (2008), is the cross-sectional area of the epilarynx tube. When the cross-sectional area of the epilarynx tube is widened and the vocal folds are firmly adducted, the source impedance is much higher than the vocal tract impedance. In this configuration, the source and filter are linearly coupled. On the other hand, when the cross-sectional area of the epilarynx tube is narrowed and vocal fold adduction levels are set to match this narrower tube, the impedances are more closely matched, thereby making the glottal flow highly dependent on acoustic pressures in the vocal tract. In this configuration, the source and filter are coupled in a nonlinear fashion. Evidence of this nonlinear coupling manifests as new frequencies created as distortion products, a lowered oscillation threshold pressure, subharmonic modulation frequencies, and sudden voice bifurcations (voice cracks) when vowel or fundamental frequency is changed.

Further evidence in support of nonlinear coupling is provided in the study by Zañartu, Mehta, Ho, Wodicka, and Hillman (2011). In this investigation, an in vivo visualization of tissue motion was implemented to inspect the effects of these instabilities. In this case study, a participant consistently exhibiting voice bifurcations during pitch glides was examined using videoendoscopy, acoustics, aerodynamics, electroglottography, and neck skin acceleration. The authors separated voice bifurcations into two categories: source-induced bifurcations and acoustically-induced bifurcations. The former, according to the authors, is primarily influenced by the tension in the thyroarytenoid muscle, while the latter is influenced by the effects of nonlinear source-filter interaction. The results of this study found three main types of instabilities in the subject's phonation: pitch jumps, pitch fluctuations, and aphonic segments. Pitch jumps (bifurcations), the main focus of this study, were the most frequent. It was found that the bifurcations caused by acoustic loading via nonlinear interactions were clearly distinguishable from those caused by muscular tension. Acoustic-induced bifurcations, caused by fundamental frequency and first formant crossings, demonstrated a pronounced visual difference in vocal fold tissue motion; whereas source-induced bifurcations, caused by changes in thyroarytenoid tension, showed a smoother transition between registers and a more symmetric behavior before and after the bifurcation. Thus, the findings by Zañartu et al. (2011) support the foundation laid by Titze (2008) with regard to the effect of nonlinear source-filter interactions.

While it may not be as relevant in the context of speech analysis, the nonlinear source-filter theory is extremely significant in singing. This is because, in singing, the shape of the vocal tract can be finely tuned to boost or alter the fundamental frequency produced by the vocal folds. In fact, Titze (2008) also emphasized the fact that entire singing styles are based on the idea that certain vowels and certain voice qualities work best with certain pitches. This phenomenon can only be explained by the effects of nonlinear source-filter coupling. Furthermore, the nonlinear interactions between source and filter are key contributors to the singer's formant, the band of frequencies causing the ringing resonant quality exhibited by classically trained operatic singers. Although the relatively new nonlinear source-filter theory has not yet received as much attention as its linear counterpart, further research of this topic will surely yield important insights and applications to the art of singing.

Chapter 5: PERCEPTION AND AESTHETICS

The discussion thus far has outlined the basic physics of sound and subsequently built upon that foundation to describe the physiological and acoustical principles governing the art of singing. Although we have examined several factors concerning the production of sound, not much attention has been given to the perception of sound. Many books on the subject of vocal pedagogy also focus disproportionately on sound production while giving little or no emphasis to sound perception. However, understanding the way in which sound is perceived is just as important as understanding the mechanics of sound production. While this is true for all musicians, it is particularly relevant for singers, because the coordinated act of singing cannot be fully consummated without the act of listening. Thus, the purpose of this chapter is threefold. First, it provides a working understanding of the anatomy and physiology of the auditory system as it relates to singing. Second, it sheds light on the topic of articulation by relating the mechanics of the auditory system to our perception of vowels and consonants. Finally, this chapter explores the topic of aesthetics in singing by examining research concerning the expression and induction of emotion through song.

The Auditory System

As mentioned in the discussion on the physics of sound, one of the necessary components for the existence of sound is a perceiver, such as the human ear. Sound perception is a sophisticated and intricate process that occurs almost instantaneously. In order to fully understand this process, it is essential to first understand the anatomy and physiology of the human auditory system. Due to the complex nature of the human ear and the limited scope of this document, the following discussion cannot be fully comprehensive; however, it will provide the foundation necessary to understanding the analytical properties of the auditory system as they relate to the art of singing. The function of the human ear is particularly complicated by the interconnectedness of its various components. Although there are numerous components, they can generally be divided into three main sections as seen in Figure 3: the outer ear, the middle ear, and the inner ear. Each section plays an important and unique role in processing the sound as it is transmitted to the auditory cortex in the brain.

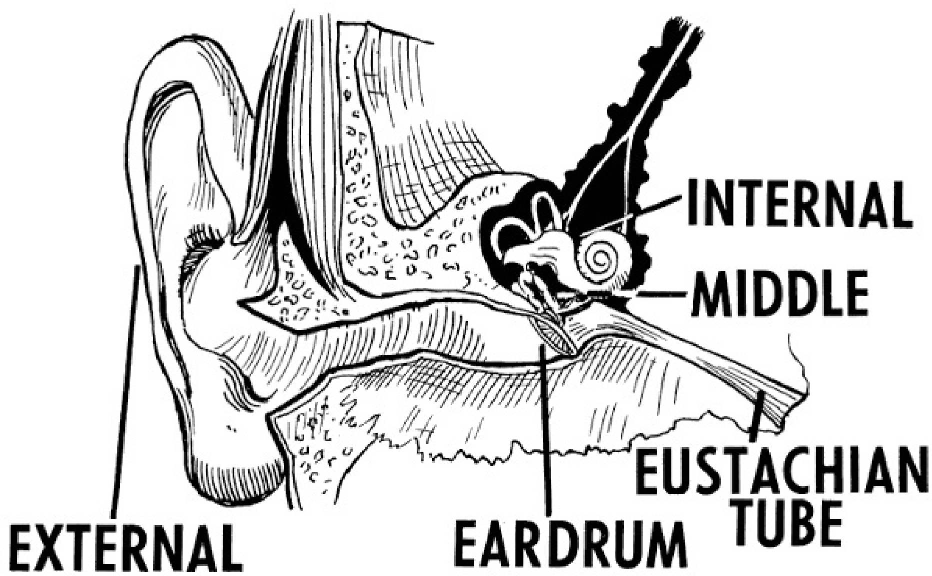

Figure 3. Components of the Ear (n.d.).
Note: In the public domain.

In order for a sound to be perceived, it must first enter through the outer ear. The outer ear consists of the visible protruding flap, known as the pinna, and the external auditory meatus, colloquially referred to as the ear canal. The unique shape of the pinna not only helps to collect the sound into the ear canal, but also to hone in on the vertical location of the incoming sound. Vertical sound localization is achieved mainly through the pinna's unique filtering properties. The pinna filters the sound in such a way that there are very slight differences in the sound spectra depending on the elevation of the sound source. These differences, which mainly affect the extremely high frequencies, are not consciously perceived as having different timbres; instead, they are perceived as having different source elevations.

Aside from assisting with vertical sound localization, a second function of the outer ear is to enhance certain frequencies. In other words, it makes the ear particularly sensitive to a certain range of frequencies. This frequency boost is primarily made possible by the unique resonance properties of the ear canal, which on average is about 2.3 cm long. As with the vocal tract, the ear canal also functions as a tube with natural resonances that filter the sound as it passes through it. Specifically, frequencies that lie within the 2000 to 5000 Hz range are amplified (Rosen & Howell, 1991, p. 237). Incidentally, this very frequency range closely corresponds to the frequency range of the formants of the vocal tract. This is extremely beneficial to vocal communication because it means that the ear is sharply tuned to distinguish the subtle formant changes that occur as a result of varying the shape of the vocal tract. These subtle formant changes, which are discussed in more detail later, are the primary cues used to distinguish between vowels. In any case, the main point to appreciate is that the resonances of the ear canal and the resonances of the vocal tract work very efficiently together to facilitate vocal communication.

Once sound has been collected by the pinna and directed through the ear canal, it impinges on the tympanic membrane, also known as the ear drum. The tympanic membrane marks the boundary between the outer ear and the middle ear. The main components of the middle ear are the tympanic membrane and three small bones referred to collectively as the ossicles. The ossicles, which are all linked to each other, connect the tympanic membrane to the main organ of the inner ear: the cochlea. When the sound wave causes the tympanic membrane to vibrate, the connected ossicles also vibrate and transmit the vibrations to the cochlea. The primary purpose of the middle ear is to serve as an impedance matching device. This is a highly important function because the medium in which the sound wave travels changes as it passes from the air-filled outer ear to the fluid-filled cochlea. When the medium of travel suddenly changes from air to fluid, as it does from the outer ear to the inner ear, much of the sound is reflected away due to the abrupt impedance mismatch. In order for sound to efficiently reach the inner ear, the middle ear must make up for this loss by matching the impedance of the outer ear to the impedance of the inner ear.

Martin and Clark (2009) mention three main mechanisms through which this is accomplished. First, the conical shape of the tympanic membrane causes it to vibrate in a complex fashion. This complex vibratory pattern of the eardrum, which essentially increases the force of vibration while reducing its velocity, assists in matching the impedances of the outer and inner ear. Second, the ossicles are positioned in such a way that it creates a mechanical lever advantage. The malleus, incus, and stapes, which are the three bones that comprise the ossicular chain, are designed in such a way that the malleus and incus have greater mass than the stapes. Consequently, the vibratory movement of the malleus and incus become more intensified at the stapes. Finally, and most importantly, the vibrating area of the tympanic membrane is much larger than the area of the oval window, which is where the stapes is coupled to the cochlea. This means that the pressure vibrations at the eardrum are focused onto the oval window, causing them to be intensified. Taken together, the complex vibratory pattern of the tympanic membrane, the mechanical lever advantage of the ossicles, and the area difference of the tympanic membrane and the oval window in the cochlea provide a total gain of approximately 30 dB. Thus, the middle ear successfully offsets the 28 dB loss resulting from the impedance mismatch when the medium changes from air to fluid (Martin & Clark, 2009, pp. 266-267).

As the sound wave is transmitted by the stapes into the cochlea, it crosses the threshold into the inner ear. The cochlea is essentially a tube that is coiled around a central pillar called the modiolus. The cochlea is often described as resembling the spiral shape of a snail's shell. The tube is divided into three distinct chambers, which can easily be seen in a cross-section. The upper and lower chambers, called the scala vestibuli and scala tympani respectively, are both filled with perilymphatic fluid. The middle chamber, known as the scala media, is filled with endolymphatic fluid (Martin & Clark, 2009, p. 303). These chambers are divided by two membranes. The scala vestibuli and the scala media are divided by Reisner's membrane, while the scala media and the scala tympani are divided by the basilar membrane.

Out of these three chambers, the most vital to sound perception is the scala media. This is because the organ of Corti, which converts the sound wave into neural impulses, is located in the scala media. The organ of Corti, which spans the entire length of the basilar membrane, contains three or four parallel rows of approximately 12,000 outer hair cells and one row of approximately 3,000 inner hair cells. Figure 4 shows a cross section of the organ of corti along with the inner and outer hair cells. Although both types of hair cells are important to the process of hearing, it is the inner hair cells that are more significant in transmitting the sound to the brain. The outer hair cells, on the other hand, play a greater role in the amplification of sound (Liberman & Gao, 2002). These hair cells are innervated by fibers of the auditory nerve (VIII cranial nerve) that extend from the modiolus into the basilar membrane (Martin & Clark, 2009, p. 304). On top of each hair cell are tiny projections known as stereocilia, which contain small tip-links that open and close with slight movement. Although the cochlea contains many other structures that have varying degrees of influence in sound perception, the structures mentioned above are the most vital and most relevant to our present discussion. With this basic anatomical knowledge of the cochlea, it is now possible to examine the way in which these structures work together to facilitate the mechanics of the cochlea.

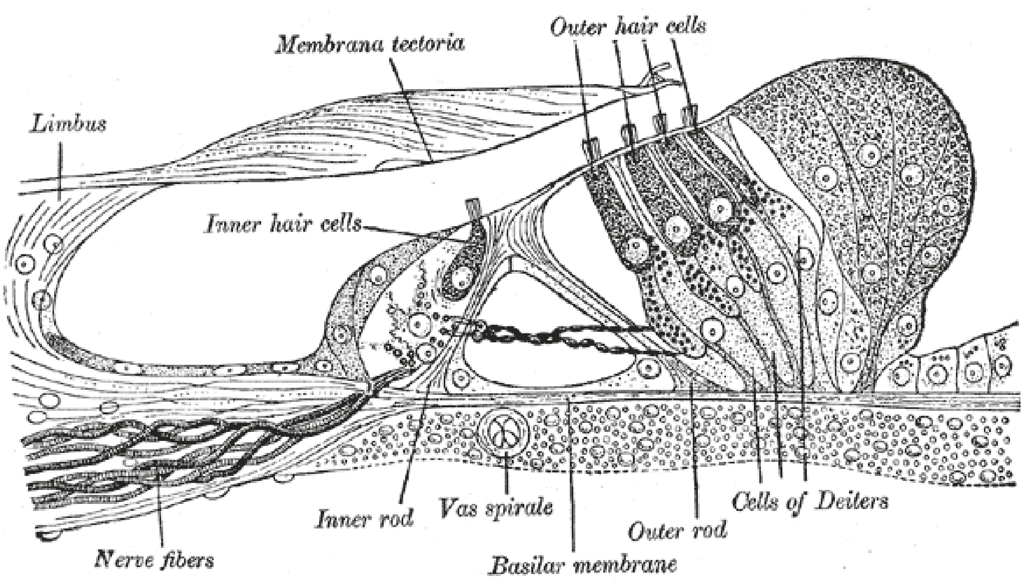

Figure 4. The Outer and Inner Hair Cells in the Organ of Corti.
Source: *Anatomy of the Human Body* (Gray et al., 1918).

In order for any sound to reach the auditory cortex, it must first be converted from a mechanical wave into neural impulses. This process, known as transduction, is the primary function of the cochlea. To outline this process, we must recall the manner in which the sound wave enters the cochlea. The stapes, the final bone in the ossicular chain, is responsible for transmitting the vibrations to the cochlea. As mentioned earlier, it is attached to a small section in the base of the cochlea known as the oval window. Thus, when the stapes is set into vibration, it displaces the cochlear fluids according to the compressions and rarefactions of the incoming sound wave. This means that the movement of the basilar membrane, which is initiated by the fluid vibrations, also corresponds to the compressions and rarefactions of the incoming sound wave. The excitation of the basilar membrane subsequently causes the stereocilia on the hair cells located in the corresponding area to be swayed back and forth. In fact, each forward and backward movement of the stereocilia also directly corresponds to each compression and rarefaction of the source signal. Due to this swaying motion, the tip-links of the stereocilia open and close, allowing the endolymph fluid to flow in and depolarize the hair cell. Depolarization of the cell immediately causes neurotransmitters to be released into the synapse, which in turn causes the nerve to fire an action potential (Rosen & Howell, 1991, pp. 246-247). Thus, it is in this way that the mechanical compression and rarefaction of the incoming sound wave is converted, or transduced, into an electrical signal that can be transmitted through the central auditory nervous system and interpreted by the brain.

Although the electrophysiology of hair cell transduction has been discussed, the mechanism by which the cochlea regulates the particular hair cells being stimulated must be mentioned. Not much was known about the mechanics of the cochlea until the pioneering work of Georg von Békésy, for which he received a Nobel Prize in 1961 (Martin & Clark, 2009, p. 308). Von Békésy found that the basilar membrane had a stiffness gradient, meaning that it is stiffer at the base of the cochlea and more compliant at the apex. When the vibrations of the stapes displace the endolymphatic fluid in the cochlea, it initiates a traveling wave in the basilar membrane. However, because the basilar membrane has a stiffness gradient, it causes the traveling wave to stop at the specific place of resonance on the basilar membrane corresponding to its frequency; higher frequencies excite the stiffer basal area, while lower frequencies excite the more compliant apical area. In other words, the cochlea is designed in such a way that the specific frequencies of the incoming sound wave excite specific places on the basilar membrane. This concept is often referred to as the tonotopic organization of the cochlea. Although this tonotopic map starts in the cochlea, this frequency-to-place mapping is preserved throughout the central auditory nervous system.

Understanding how this tonotopic organization assists in pitch perception is of particular interest to all musicians, including singers; hence it will be discussed briefly now. When the basilar membrane is excited by a simple sinusoidal wave, the peak excitation occurs at a single point on the basilar membrane. However, this is not the case with complex sounds. As mentioned earlier, complex sounds are comprised of more than one sine wave. Thus, when the basilar membrane is excited by a complex sound, each of the simple sine waves excite a specific area on the basilar membrane. In other words, the basilar membrane essentially performs a Fourier analysis, breaking down the complex sound into its simple constituents and sending each signal to the brain through the central auditory nervous system. It is only when the signals reach the auditory cortex that the separate signals are reintegrated and interpreted as a whole.

Before the sound wave can be interpreted by the auditory cortex, it must pass through six subcortical nuclei (cochlear nucleus, trapezoid body, superior olive, lateral lemniscus, inferior colliculus, and the medial geniculate body), where it is thoroughly analyzed (Martin & Clark, 2009, pp. 341-342). Although an in-depth understanding of the physiology of each of the subcortical nuclei is probably unnecessary and of little use to singers, there are two important concepts regarding central auditory processing that should be highlighted. First, it should be emphasized that the tonotopic organization found within the cochlea is preserved throughout each of the subcortical nuclei. Furthermore, the tonotopic maps found in the central auditory structures, unlike in the cochlea, are represented in three dimensions. In the cochlea, frequency is mapped to the basilar membrane two-dimensionally; points closer to the base of the cochlea are tuned to high frequencies, and points closer to the apex are tuned to low frequencies. However, in the central auditory structures, frequencies are mapped to three-dimensional slabs called isofrequency laminae. The term isofrequency is used because all areas on a single lamina are tuned to the same frequency. Each isofrequency lamina is comprised of a mixture of different cell types, each with different processing capabilities. In this way, each frequency is processed by various classes of cells, thereby modifying the input to the brain.

This leads to the second important concept involved in central auditory processing. Each of the subcortical nuclei have unique analytic strengths due to the concentration of cell types within each structure. For example, the superior olivary complex excels in determining the horizontal location of the sound source by analyzing extremely minute differences in time and intensity between the inputs of each ear (Martin

& Clark, 2009, p. 342). Essentially, if sound arrives more quickly or loudly on one side, then the sound source must be located on that side. On the other hand, the dorsal cochlear nucleus, which has a large concentration of fusiform cells, excels in deciphering cues for vertical sound localization. The fusiform cells are extremely sensitive to the fine structural changes in the high frequency spectrum resulting from the filtering effect of the outer ear. In this way, each subcortical nucleus is sensitive to different auditory parameters, which in turn provide the auditory cortex with sufficient data for an accurate interpretation of the signal.

Articulation and Perception

Sound perception is an intricate and sophisticated process that involves the resonance of the outer ear, the amplification of the middle ear, and the transduction and tonotopic organization of the inner ear, as well as the neural analysis of the subcortical nuclei. Nevertheless, understanding this process is of great importance to singers. It must be heavily emphasized that sound is dualistic by nature. As mentioned earlier, the coordinated act of singing cannot be fully consummated without the act of listening. One aspect of singing that can be better understood with the knowledge of sound perception is articulation. In the previous chapter, the relationship between vocal formants and vowel relationships was briefly mentioned. Now that the fundamentals of sound perception have been laid out, it is possible to discuss in greater detail how changing vocal formants influence the perception of vowels and consonants.

As mentioned earlier, vocal formants are peaks in resonance determined by the shape of the vocal tract. Thus, by changing the shape of the vocal tract, the frequency areas of the formants are altered. The primary mechanism of changing vocal tract shape is through the movement of the tongue. In fact, tongue movement is so important that vowels are classified according to the position of the tongue; high and low vowels refer to the vertical height of the tongue, while front and back vowels refer to the horizontal advancement of the tongue (Ferrand, 2007, pp. 205-206). The height of the tongue is correlated with the first formant frequency (F_1) of the vocal tract, and the advancement of the tongue is correlated with the second formant frequency (F_2). Although the precise frequencies of F_1 and F_2 have some degree of variability within each individual voice, the two general rules are that F_1 is lower in high vowels and higher in low vowels, while F_2 is lower in back vowels and higher in front vowels (Ferrand, 2007, pp. 205-206). Thus, it is the relative distance between F_1 and F_2 that provide the primary cue for vowel identification.

Consonants, unlike vowels, are produced much more rapidly by various combinations of articulators. There are numerous articulators that are employed in the production of consonants which include the tongue, teeth, alveolar ridge, hard and soft palates, lips, and sometimes even the velum. Due to the variety of consonant sounds, the acoustic cues involved in consonant identification are more complex than those used for vowel identification. Aside from the location of the formant frequencies, other cues used to identify and distinguish between consonants include the fundamental frequency of the consonant, the voice onset time, as well as the duration of the consonant sound (Ferrand, 2007, p. 267). Since there are a variety of consonants, the relevance of each cue depends upon the type of consonants being distinguished. For example, one of the primary cues used to distinguish voiced stop consonants such as [b] and [d] from unvoiced stop consonants such as [p] and [t] is the voice onset time, which is the amount of time it takes for vocalization to occur after the silent articulatory stop. In unvoiced stop consonants, the voice onset time takes longer than for voiced stop consonants. In fact, in voiced stop consonants, the onset of the voice may sometimes even occur before the articulatory stop (Ferrand, 2007, p. 217).

Although the identification of consonant and vowel sounds involve different acoustic cues, they both share a common characteristic that proves extremely helpful in understanding speech. Aside from sibilants and fricatives, most common speech sounds, both consonants and vowels, lie within the frequency area of 500-4000 Hz (Ferrand, 2007, p. 273). The auditory system, which is most sensitive at these very frequencies, thereby facilitates the processing of speech sounds. This sensitivity is due to several reasons: the resonance of the outer ear canal, the amplification of the middle ear, as well as inner ear amplification provided by the movement of the outer hair cells. Thus, while the range of human hearing extends from 20-20,000 Hz, the most sensitive range of frequencies is from 500-4000 Hz (Appleman, 1974, p. 144). Due to this sensitivity, it is possible to perceive slight differences in the formant frequencies, the time of vocal onset and various other minute cues, which otherwise may go undetected. Sibilant and fricative consonants are the only outliers, as they usually lie above 4000 Hz, which is higher in frequency than the majority of speech sounds. Thus, individuals with high frequency sensorineural hearing loss resulting from age-related changes, noise exposure, harmful side effects of medications, or a variety of other reasons, generally experience difficulty distinguishing sibilant sounds such as [s] from fricative sounds such as [f]. In fact, with even greater loss in the high frequencies, some may have difficulty distinguishing other consonant sounds as well, as consonant cues tend to be higher in frequency than vowel cues.

Articulation not only facilitates communication by providing intelligibility, but it also serves as a vehicle for conveying emotion. The manner in which a word or sentence is articulated can communicate and induce a variety of emotions. Furthermore, the way in which it is perceived can also influence the emotions being conveyed. Finally, during the process of singing, articulation occurs within the context of music. Thus, emotions being conveyed through song are further affected by their musical context. In order to fully understand how emotion is conveyed and elicited through the art of singing, we must first consider emotion as it relates to music as well as emotion as it relates to speech. Synthesizing these aspects of emotion will allow for a better understanding of emotion as it relates to the art of singing.

Emotion and Aesthetics

Emotion in music is a well-researched topic in the field of music psychology. There have been numerous studies that bring to light various relationships between emotion and music. Still, it is evident from the critiques of these studies that the subject, though well investigated, is still somewhat shrouded in mystery and there is much more work to be done in the field. Research in this field can be divided into two distinct categories: those done through a cognitivist perspective, and those done through an emotivist perspective. The cognitivists are more concerned with the expression and communication of emotion, while the emotivists are more concerned with the induction of emotion. In other words, cognitivist research focuses on how emotion can be understood and decoded by the brain—cognition of the emotion. On the other hand, emotivist research prioritizes how emotional feeling is elicited and induced—feeling the emotion. It must be emphasized, however, that dichotomizing these two perspectives does not imply that they are mutually exclusive. On the contrary, both perspectives are essential in order to gain a comprehensive understanding of emotion in music.

Inclined more towards a cognitivist approach, Timmers and Ashley (2007) found that ornamentation can be successfully utilized to communicate basic emotions through the variance of certain acoustical elements such as loudness, duration, timing, complexity, and density. A similar experiment was conducted by Patrick Juslin (1997). This experiment consisted of guitarists attempting to musically express four basic emotions in the same musical passage. The results for this study not only agreed with the aforementioned study by Timmers and Ashley, but it also revealed that the individual musical elements are all interrelated. In a later article, Juslin (2003b), taking on a true cognitivist approach, designed the GERMS model, which illustrates that expression in musical performance is comprised of five main components: ***G***enerative rules, ***E***xpression of emotional intent, ***R***andomness due to human limitations, ***M***ovement patterns, and ***S***tylistic novelty. Generative rules refer to the composer's intent, which is generally revealed in the musical score. Expression of emotional intent, on the other hand, is dependent on the performer's intent as well as the decoding ability of the audience; the performer utilizes auditory cues to convey emotion in a passage. Randomness refers to variations due to human imperfection; this can include such elements as coughs, mistakes, and timing incongruences. The term movement patterns simply indicates musical elements that can be related to biological patterns of motion; for instance, an accelerando can be likened to a runner beginning to pick up speed. Each of these five facets, according to Juslin (2003b), plays a pivotal role in the conveyance of emotion in musical performance.

On the emotivist side, researchers were far more concerned with the emotion-inducing properties of music than the communication of emotion through music. In his article, “Does Music Induce Emotion? A Theoretical and Methodological Analysis,” Vladimir Konečni (2008) examined the existing model of the direct induction of emotion through music. In his examination, he brought to light several of the problems in the prior studies supporting the model. His criticisms, beginning with the oversimplification of the cognitivist versus emotivist dichotomy, encompass the linguistic ambiguities and the methodological ambiguities present in many of the prior studies. In order to clarify one such ambiguity, he distinguishes the difference between the expression of emotion and the induction of emotion by demonstrating conceptually that induction is only one of the possible reactions to the expression of emotion; in other words, the emotion being expressed need not necessarily be experienced by the perceiver. Furthermore, he addresses how this issue can become a methodological pitfall by explaining that subjects, when asked to rate their emotional response to music, have a tendency to take into account the music’s expressive qualities and cause the ratings to be inflated.

In the same article, Konečni inspects those prior studies linking music to the direct induction of basic emotions (e.g., happiness, sadness) using his Prototypical Emotion-Episode Model (PEEM) as a tool for analysis. Utilizing this model, he discovers that due to the inherent issues in the prior studies—such as imprecise wording, methodological issues, and the heavy impact of associations on the emotional capacity of music—there are, in reality, very few genuinely induced emotions resulting purely from music. Three “musical emotions,” are proposed by Konečni as a replacement for the basic emotions previously thought to be induced directly by music. The three “musical emotions” include, in order of increasing rarity: thrills and chills, being moved, and aesthetic awe. These emotions do not fit into Konečni’s PEEM (prototypical emotion-episode model) as regular emotions, which consequently led some researchers to question why they should be labeled “emotions” in the first place. Due to this dissatisfaction with the terminology, Konečni proposed the Aesthetic Trinity Theory, which “postulates the singular significance in the arts of the related states of aesthetic awe, being moved, and thrills or chills” (p. 123). Since this theory was only recently developed, it has yet to be extensively studied or tested by other researchers. Nevertheless, it still proves useful in conceptualizing the effect of music on the induction of emotion.

The human voice is one of the fundamental means of conveying emotion. In the field of speech-emotion analysis, most research is concerned with exploring the chief characteristics of the human voice in the transmission and perception of emotion. Klaus J. Scherer, one of the leading researchers of vocal emotion, has extensively studied the characteristics and influence of emotion as it relates to speech. Scherer (1995) revealed that there are relatively few studies focused primarily on emotion in the human voice; instead he states that research correlating facial expression and emotion far outnumber the former. He maintains this statement in a later article, thus reinforcing the need for study in the field (Scherer, 2003). The component process theory, which was proposed by Scherer (1986), is still the only existing theoretical model to specifically account for emotion in speech. This model defines emotions as the synchronization of many different cognitive and physiological components; cognitive, neurophysiological, motivational, expressive, and subjective-feeling components. The model asserts that the changes in the components are all multiply interrelated, thus implying that emotion is a process (Scherer, 1987).

Scherer also briefly attempted to discuss the role of emotion in singing. He drew on an idea presented by Helmholtz—that music was thought to be discovered from the endeavor to imitate the human voice—and argued that if this is the case, then naturally it follows that musical emotion would be most effectively conveyed through the human voice (Scherer, 1995). He went on to affirm that the emotional characteristics correlated with speech are also correlated with singing. Still, he argued that the art of conveying emotion through singing depends upon three pivotal factors. The first, according to Scherer, is the composer's emotional intentions conveyed through the score. The second is the singer's interpretation of these emotions being expressed. Finally, the singer's own psychological emotional state also plays a key role in the expression of emotions in singing (Scherer, 1995). Although Scherer noted the lack of studies involving emotion and expression in singing that is not to say that there have not been any in the field.

Some of the studies specifically investigating emotion in singing include one by Mayumi Adachi and Sandra E. Trehub (2000) investigating the relationships between age, gender, and the ability to encode and decode both visual and auditory emotions present in children's singing; and the study by Albert LeBlanc and Carolyn Sherrill (1986) measuring the influence of high and low vibrato as well as the sex of the performer upon listener preference. Although these studies do not provide much information regarding specific emotions and are difficult to apply to classical singing, they do provide evidence that acoustic vocal parameters, such as vibrato, impact the perception of the music, thereby influencing the perception of the emotion as well. A more relevant study to the emotional characteristics of singing was undertaken by Livingstone, Thompson, and Russo (2009). In this study, it was found that certain facial features, such as the corners of the lips and the eyebrows, play a pivotal role in visually conveying emotions during the act of singing.

One of the most comprehensive investigations on this topic was Juslin's (2003a) grand undertaking in which he examined 104 studies on vocal expression and 41 studies on music performance attempting to track cross-modal similarities and patterns. The study reviewed "parallels between vocal expression and music performance regarding the accuracy with which emotions were communicated" as well as the "emotion-specific patterns of acoustic cues used to communicate each emotion" (p. 770). The results of the study indicated that there were indeed cross-modal similarities between the two channels, and Juslin argued that these similarities demonstrated that vocal expression was in fact the model for musical expression rather than the reverse. In the end, Juslin, in agreement with Scherer, stated that this topic still has much to be studied and also provided various suggestions for the clarity and effectiveness of future research (Juslin, 2003a).

At this point in our discussion, we have established a solid scientific foundation of the various acoustical elements involved in the art of singing. Beginning with the physical properties of sound, this discussion examined the physiological aspects of breathing, phonation, registration, resonance, articulation, perception, and emotional aesthetics through modern scientific research. With this solid foundation, it is now possible to explore the scientific principles underlying the major tenets of *bel canto*. In doing so, this project strives to begin bridging the gap between the scientifically-oriented schools and the traditionally-oriented schools of singing.

Chapter 6: Bel Canto: Art, Science, and Legacy

As mentioned at the outset, the overarching purpose of this document is to illuminate the acoustic principles that are inherent within the teachings of *bel canto*. Throughout the previous chapters, the reader was provided with a working understanding of the mechanics of the singing voice and its peripheral fields. With this scientific groundwork in place, it is possible to more closely analyze specific *bel canto* principles and apply them to our current scientific understanding of the vocal mechanism.

The growing schism between science and tradition in the subject of voice instruction has become a heated point of dispute among teachers, pedagogues, and students alike. Both sides have laid out several arguments against each other. For example, Cornelius Reid, a well-known authority on *bel canto*, was of the opinion that voice scientists, who according to him, are often in sharp disagreement with some of the most fundamental teachings, have yet to contribute significantly to the advancement of the art of singing (Reid, 1950, p. 3). On the other hand, voice scientists maintain that traditional teaching methods that use vague and inexact terminology are susceptible to misunderstandings over time. Richard Miller, a well-known voice scientist and pedagogue, stated that technical knowledge is inseparable from the art of singing and using precise scientific language is essential to communicate concrete concepts (Miller, 1996, pp. xv-xvi). However, despite the seemingly incompatible nature of these two perspectives, it must be clearly emphasized that science and tradition are not mutually exclusive. Even Reid admits that the traditional teachings of the *bel canto* style and scientific principles *cannot* be at odds with each other:

> Because of a complete absence of scientific knowledge upon subjects related to voice it was impossible for the early systems of training to be founded upon principles other than those growing out of empirical observation. Yet, in view of the extraordinary accomplishments of the vocalists trained by these procedures the instruction must have been securely based upon principles that were both scientifically sound and aesthetically pleasing. (Reid, 1950, p. 15)

From this statement it is very evident that, despite having originated through traditional empirical observations, the age-old teachings of *bel canto* were undoubtedly supported by scientific principles. Thus, the aim of this chapter, then, is to highlight the scientific evidence present within *bel canto* teachings by synthesizing the scientific foundations laid out in the previous chapters with the major precepts of *bel canto*. In doing so, this project helps contribute toward the noble goal of bridging the gap between science and art in vocal instruction.

Bel Canto Literature

Although the anatomical and physiological aspects of singing were laid out in great detail in earlier chapters, these aspects have not been examined with regard to the *bel canto* literature. However, before delving into the literature, it must be re-emphasized that *bel canto* was primarily an oral tradition passed down from teacher to student organically. As such, the written pedagogical literature describing these methods are scarce. Even esteemed vocal pedagogue, Richard Miller, asserted that a specific codified system of *bel canto* is impossible to assemble and that the term itself has become "a twentieth-century shibboleth, with opposing methodologies staking out highly suspect claims for its possession" (Miller, 1996, pp. xx-xxi).

Despite Miller's rather blunt criticism of the modern state of voice teaching and the lack of a written codified system of *bel canto*, it is not to say that there are no useful treatises or manuals from that era. On the contrary, as mentioned in the beginning of this document, there are multiple sources that provide excellent insight into the philosophies, principles, and practices of the early systems of singing. In fact, some authors have even compiled and analyzed the writings of several prominent teachers and historical figures of the *bel canto* era. These works not only provide greater accessibility to these historical writings, but they also paint a comprehensive image of the beliefs and practices of the time. The works of Cornelius Reid, Philip Duey, and James Stark all provide a bird's-eye view of *bel canto* by comparing and contrasting the writings of influential figures of the era such as Pier Francessco Tosi (1653-1732), Giambattista Mancini (1714-1800), Francesco Antonio Pistocchi (1659-1726), and many others. In addition to examining these comprehensive secondary sources, we will also review some primary sources, such as direct writings and treatises from the era.

Breathing

Although breathing had always been viewed as important to the art of singing, Caccini was probably the first to articulate so well the necessity of breath management to vocal artistry. In his landmark treatise, *Le Nuove Musiche*, Caccini stressed the importance of breathing to his noble and nonchalant manner of singing which he refers to as "una certa nobile sprezzatura di canto" (Stark, 1999, p. 161). However, he advised that a cautious management of the breath is extremely vital to this noble manner of singing and that the singer must be careful not to spend too much breath on the various vocal inflections utilized in this style so that it will not fail him when it is most needed (Duey, 1980, p. 74). The main inflection that Caccini was alluding to in this passage was the *messa di voce*, a technique in which the singer, while sustaining a single note, swells and diminishes the voice in a way that displays the entire dynamic range. Although Caccini was the first to describe it, the *messa di voce* remained a core practice of *bel canto* singing throughout the years and has been discussed a great deal in the literature. Giambattista Mancini, for example, also stressed that the acquisition of proper breath management was an essential prerequisite to successfully executing the *messa di voce* (Duey, 1980, p. 76).

Similar to these views, much of the sentiment expressed during the early years of *bel canto* regarding breathing was relatively simple and straightforward. Most of these teachings, as with Caccini's teachings, merely involved quiet inspiration and proper management of the breath according to the musical phrase (Stark, 1999, p. 93). However, the subject of breathing began to receive increased attention as it began to be described with greater detail during the 18th and 19th centuries. Both the Garcia and Lamperti schools had quite definite views regarding breathing, and each offered distinct guidelines pertaining to its methodology.

The Lamperti school emphasized breathing quite a bit. In his famous treatise, *The Art of Singing*, Francesco Lamperti espoused the views of Dr. Louis Mandl, a famous French physiologist who was frequently quoted by the Lampertis in several of their manuals, and described three conventional types of breathing:

> Respiration is of three types—the Diaphragmatic or Abdominal, the Lateral, and the Clavicular. In abdominal respiration, the diaphragm, a large muscle at the base of the thorax, is the motive agent; its action is shown by the movement of the stomach. Lateral respiration is caused by the dilation of the middle part of the thorax, and is shown by the movement of the ribs and breastbone. Clavicular respiration is caused by the dilation of the upper part of the thorax, and is shown by the movement of the breastbone, ribs, collar-bones, shoulders, vertebrae, and sometimes also the head. These different types of respiration are often combined, or rather succeed one another; for instance, a continued abdominal respiration will become also lateral. (F. Lamperti & Griffith, 1890, p. 20)

From this description, it is clear that Lamperti believed that breathing involves a combination of several components. Later on in the same chapter, he warns singers of the unwanted tension that clavicular breathing brings about in the laryngeal muscles. Although Lamperti frowned upon clavicular breathing, he did not disapprove of either abdominal or lateral types of breathing. This is particularly evident in the following passage regarding breathing in female singers:

> It is a mistake to suppose that the clavicular type of breathing is natural to women. . . The pressure of the corset upon the abdomen or in some cases the abnormal development of the stomach not permitting of the natural descent of the diaphragm, the respiration becomes lateral; the movement of the ribs and breast-bone causes the rise and fall of the bosom, thus leading one to believe that the breathing is clavicular; but will be seen that there is no movement of the collar-bones, and so we may be sure that the natural type of respiration in the case of women, if not abdominal, is lateral. (F. Lamperti & Griffith,1890, pp. 20-21)

Thus, it is evident that Lamperti prefers both lateral and abdominal breathing to clavicular breathing. However, some scholars assert that a closer look at his writings reveals a slight preference towards abdominal or diaphragmatic breathing (Stark, 1999, p. 100). This interpretation of Lamperti's views is perhaps due to the fact that abdominal breathing is the least likely of the three types to interfere with vocal production as the inspiratory muscular movement of the diaphragm occurs furthest away from the vocal mechanism. However, it must be kept in mind that Lamperti also believed that abdominal and lateral respiration were often executed in a combined fashion, as mentioned in the earlier excerpt. Thus, it is safe to conclude that while Lamperti did indeed give special emphasis to deep breathing, he did not imply that the lateral expansion of the ribs was to be avoided altogether.

Francesco Lamperti's son, Giovanni Battista Lamperti, also believed in a combined style of breathing. His instruction, transcribed by his student William Earl Brown, in his book *Vocal Wisdom* is evidence of this perspective:

> It is a mistake to breathe in just one part of the body . . . When the top and bottom of the lungs are equally full of compressed air, the voice will focus in the head, and awake all the resonance in the head, mouth, and chest. (G. B. Lamperti & Brown, 1957, p. 43).

This sentiment makes it apparent that the Lampertis believed that the lateral expansion of the ribcage should be combined with the downward descent of the diaphragm to obtain the most efficient means of inspiration. It is also important to note that in this passage, Lamperti also mentioned the expression "compressed air" (sometimes also referred to as compressed breath in his writings). This concept, which was meant to refer to the steady subglottal pressure of the inspired breath, was critical to the Lamperti school of singing and was mentioned repeatedly in their writings.

The Lampertis were also responsible for introducing and popularizing the concept of *appoggio*. This term must be mentioned here briefly as it is an extremely prominent term that appears a great deal in the *bel canto* literature as well as in the general vocal pedagogical literature. Through the years, the term has been used to imply a variety of meanings. In its basic essence, *appoggio*, which is derived from the word *appoggiare* meaning "to lean," refers to the balance between the opposing forces of inspiratory and expiratory muscles. This balance essentially creates a strong foundation or support for the voice. However, the Lampertis used the term in a broader sense to apply not only to the balance between inspiratory and expiratory muscles, but also to an array of other factors including vocal onset, glottal closure, vocal tract positon, airflow and breath pressure, legato, the *messa di voce*, as well as good intonation (Stark, 1999, p. 102). To the

Lampertis, *appoggio* was not merely a single technique; instead it was viewed as a complete system of breathing.

Although the Garcia school was often perceived as being at odds with the Lamperti school, Garcia II also more or less agreed with the Lampertis on the issue of breathing. While Garcia spoke only briefly about breathing in his treatise, what he said was quite important. He also emphasized that the combination of diaphragmatic and lateral breathing allows for the most efficient use of lung capacity:

> To ensure easy inspiration, it is requisite that the head be erect, the shoulders thrown back without stiffness, and the chest expanded. The diaphragm should be lowered without any jerk, and the chest regularly and slowly raised. This double movement enlarges the compass or circumference of the lungs; first, at their base, and subsequently throughout their whole extent, leaving them full liberty to expand, until they are completely filled with air. (Garcia & Garcia, 1924, p. 6)

Thus, Garcia observed the improved efficiency of combining the lateral and diaphragmatic types of breathing. In addition, this passage makes it clear that in order to best facilitate this breathing, the chest should be elevated and the diaphragm lowered. This posture was asserted by other prominent teachers of singing as well. In fact, even before Garcia, Mancini also noticed that an elevated chest and a well-developed thorax were essentials to good singing (Duey, 1980, p. 75).

Although there were several different opinions and observations regarding breathing for singing as seen through the writings of the prominent teachers of the era, it must be kept in mind that each emphasized a various aspect of the mechanism and that their ideals may not necessarily be mutually exclusive. For example, one of Garcia’s students, Hermann Klein, advocated in his own book the importance of a high chest, diaphragmatic breathing, and compressed breath. These ideals, which reflect the teachings of both the Garcia and Lamperti schools, make it clear that their principles were not mutually exclusive of one another. In fact, these principles are all supported by modern scientific findings.

The physiology of the combined downward movement of the diaphragm and the lateral movement of the ribcage was described earlier in this document. It was established that the downward motion of the diaphragm allows the lungs to expand downward vertically, while the lateral motion of the ribcage through the contraction of the external intercostal muscles allows the lungs to expand laterally and upward. Thus, it is evident that a coordinated movement of these two processes allows the lungs to expand the most without causing any unwanted tension in the larynx.

Furthermore, the concept of compressed breath is also scientifically sound. Although it is referred to as subglottal pressure in modern scientific literature, compressed breath is necessary to properly sustain speech and singing. As mentioned earlier, the role of subglottal pressure is to maintain the constant airflow needed to actuate and sustain vocal fold motion. This subglottal pressure is created and stabilized by the coordinated use of the expiratory and inspiratory muscles. In other words, the opposing forces of the inspiratory muscles and expiratory muscles create a steady subglottal breath pressure that provides a stable support for vocal fold oscillation. This concept may have been at the heart of *appoggio*, observed and described by the Lampertis.

Finally, the slightly elevated chest position advocated by Mancini, Garcia, Klein, and many others has also found support in modern science. It was mentioned earlier that a study by Hoit and colleagues (1996) found that one major difference between the breathing process in classically trained singers and untrained singers is the elevated chest and tucked in abdomen immediately following inspiration. This posture allows the singer to more efficiently utilize the breath, especially in long or difficult phrases. As mentioned earlier, it increases efficiency in two important ways. First, the expansion and elevation of the ribcage places it at a mechanical advantage for making quick expiratory pressure adjustments. This is because it works in conjunction with gravity to make sudden increases in expiratory pressure. Second, the tucking in of the abdomen places the diaphragm at a mechanical advantage for making quick inspiratory pressure adjustments. This is mainly because the recoil forces of the abdominal muscles work jointly with the expansion of the diaphragm.

Thus, it is evident that these *bel canto* teachings regarding breathing are indeed scientifically sound. Although many these observations may not have initially been discovered through scientific means, it is now possible to uncover the scientific principles within them.

Vocal Production

The production of the voice is comprised of two major components: the vibration of the vocal folds and the adjustments made by the vocal tract. Taken together, these two components comprise the source and filter of the vocal mechanism. As established earlier in this document, the characteristics of the source determine the pitch of the sung note, while the characteristics of the filter determine its quality or timbre. Although these two terms were not mentioned specifically by teachers of the time, many of them did indeed observe and classify these concepts in their own way.

Early views on voice production were fairly simple. Many teachers of singing merely emphasized a free and natural vocal production. One of the main concerns of the early teachers was the degree of openness of the mouth when singing. Both Tosi and Mancini cautioned that opening the mouth too much or too little could result in the obstruction of a natural and free tone (Duey, 1980, p. 102). Aside from this idea of free tone production, there was not much emphasis on vocal resonance or projection in the early treatises. Despite the scarce descriptions of resonance mechanics in the early literature, some of the observations made by teachers of the time suggest that resonance was indeed an important and desirable component of good singing. Mancini in particular observed the relationship between the shape of the mouth and vocal resonance:

> Experience proves that the opening of the mouth is what directs and regulates the voice. In fact, the resounding quality of the voice always depends upon the shaping of the position of the mouth, when there is the natural strength of the chest and a harmonious disposition of the vocal organs. (Mancini & Buzzi, 1912, p. 92)

Mancini noticed that the shape of the vocal tract had a significant and direct impact on the quality and resonance of the voice. Although he did not articulate how or why this happens, it is clear that he understood, at some level, the filtering effects of the vocal tract on the vocal source. It could even be argued that, in essence, what he described in the above passage is an extremely simplified version of the source-filter model.

Vocal resonance, although not addressed overtly in the early literature, was undoubtedly an integral aesthetic ideal of *bel canto*. Throughout the years, the tonal preferences of the voice were documented using a variety of terms and in several languages. One of the key terms pertaining to vocal timbre that gained widespread fame in the Italian school of singing was *chiaroscuro*. Although it was an integral concept advocated in the Lamperti school, the development of this term did not occur overnight from any one particular teacher or scholar of the voice. Instead, it developed gradually through the endeavors of many teachers and scholars as they transformed their tonal ideals into a pedagogical concept (Stark, 1999, p. 34). The term *chiaroscuro*, which is derived from the combination of two Italian words *chiaro*, meaning light, and *oscuro*, meaning dark, was originally used in describing drawings and paintings of the Renaissance. In essence, the term refers to the balance of two contrasting yet complementary aspects of the tone: brightness and darkness. The brightness stems from the rich overtones produced by the laryngeal source (i.e., vibration of the vocal folds) while the darkness of the tone is determined by the filtering effects of the vocal tract (Stark, 1999, p. 34). Thus, upon closer inspection, it is evident the term *chiaroscuro* developed as an endeavor to describe the tone quality resulting from the ideal interaction between source and filter.

The initial source-filter theory, developed through the pioneering work of Johannes Muller (Muller & Baly, 1848) and further developed by Gunnar Fant (1960), viewed the interaction between the source and filter as linear. This means that, although the source and filter both contribute to the output sound, neither one can behave in a way that changes or influences the other. In this model, the physical properties of the vocal tract cannot change or alter the vibratory characteristics of the vocal folds. While this linear model has been used as the standard model of speech analysis throughout the years, relatively recent findings, mentioned earlier in the document, indicate that there are indeed some non-linear interactions between source and filter.

The concept of a vocal source and filter was also observed and described by Manuel Garcia II. In fact, according to Stark (1999), it was Garcia who first systematically explained vocal production in this twofold manner (p. 36). Garcia details this concept in his *Traité*:

> The moment that a sound is emitted, it becomes subject to the influence of the vocal tube through which it passes; which tube, having the power of lengthening or shortening, contracting or expanding, and of changing its curvilinear form to that of a right angle, most perfectly fulfils the function of a reflector to the voice. . . . We shall understand these movements if we consider the vocal tube as a deep and highly-elastic pipe, beginning below at the larynx, forming a curve at the arch of the palate, and ending above at the mouth;—a tube, which, when at its shortest dimensions, forms only a slight curve, and at its longest, nearly a right angle; . . . The short and gently curved shape produces the clear timbre, while the sombre is caused by the lengthened and strongly curved form. (Garcia & Garcia, 1912, p. 4)

Garcia's portrayal of the vocal mechanism clearly demonstrates the filtering effects of the vocal tract in creating the various timbres of the voice. The two timbres described in the passage above, clear and sombre, also reflect the bright and dark elements of *chiaroscuro*. Garcia also emphasized numerous times that the these two opposing timbres, clear and dark, could borrow from each other in order to produce a wide array of vocal colors (Stark, 1999, p. 39).

Although the nonlinear interactions between source and filter were not mentioned explicitly by *bel canto* teachers, the principles that they advocated were specifically predicated on this relationship. One such principle is the concept of *aggiustamento*, or vowel modification, which was advocated by many *bel canto* teachers, including Garcia:

> But, in order to attain evenness of voice, a singer should, by clever management, modify a vowel, insensibly rounding it as the voice ascends, and brightening as it descends; by this means, a seeming equality results from a real, but well-concealed inequality of the vocal sound. This precept applies to each register throughout the entire compass. If a vowel remained constantly open, as the *a* when sounded in the word father, it would give brightness to the low and middle sounds, while high notes would be shrill and shriek; whereas a vowel that is invariably covered, like the o in the word note would give richness to high notes, and make low ones veiled and dull. (Garcia & Garcia, 1912, p.43)

This concept of *aggiustamento* (i.e., vowel modification) hinges on the notion that certain vowel shapes are better suited to certain pitches. Thus, it is quite clear that these source-filter interactions were certainly observed and recognized, on some level, by *bel canto* teachers.

Another key aspect of voice production that received great attention among the early teachers is registration. Because many different views have sprung up, especially in recent years, regarding the nature and number of registers in the human voice, it is necessary to discuss here the subject of vocal registers as it was understood and taught by the key figures of the *bel canto* era. There was not much controversy regarding the subject of registers among the early teachers. Historically, the old Italian school maintained that there were only two registers: the chest and head registers (Stark, 1999, p. 57). This notion can be seen in several of the early vocal treatises. Caccini clearly delineated two registers in his *Le Nuove Musiche* and furthermore showed preference for the full natural voice in contrast to the weaker falsetto:

> In this connection it will be useful to note that when he who professes this art is to sing alone to the archlute or some other stringed instrument without being constrained to accommodate himself to others, let him choose a key in which he can sing with a full, natural voice, avoiding falsetto. . . From the falsetto voice no nobility of good singing can arise; that comes from a natural voice, comfortable through the whole range, able to be controlled at will. . . (Caccini & Hitchcock, 1970, p. 56)

Apart from Caccini, several others also subscribed to the two-register theory. Mancini, for example, also clearly wrote in favor of this model in his treatise:

> The voice ordinarily divides itself into two registers, one called chest register and the other head register, or falsetto . . . Every student, whether he is soprano, contralto, bass or tenor, can easily know the difference between these registers . . . The great art of the singer consists in acquiring the ability to render imperceptible to the ear, the passing from the one register to the other. In other words, to unite the two, so as to have perfect quality of voice throughout the whole range, each

> tone being on a level with your best and purest tone. (Mancini & Buzzi, 1912, pp. 58-59)

This two register model was undoubtedly the accepted standard during the early years of *bel canto*. However, towards the end of the 18th century, a new three-register model was proposed by some authors. This three-register model, which was primarily utilized to describe female voices, gave rise to several new register terms, which only added further confusion to the topic. Terms such as *falsetto, voce di testa, voce finta, voce di mezzo petto, voix mixte, voix du gozier, and voix sombree ou couverte* were all used to refer to registers above the normal *voce di petto* or chest voice (Stark, 1999, p. 68).

During Garcia's time, the topic had become quite complicated. Although the two-register theory was still widely accepted as the standard, the newer three-register perspective was gaining popularity. Thus, Garcia began his theory by establishing a definition for the term register:

> The word register means, a series of consecutive and homogeneous sounds produced by the same mechanism, and differing essentially from other sounds originating in mechanical means of a different kind. (Garcia & Garcia, 1912, p. 4)

Garcia's theory of registers was somewhere in between the two, creating a sort of bridge between them. He initially acknowledged only two main registers, which he called the chest register and the falsetto-head register (Stark, 1999, p. 68). He described the latter as a single register consisting of two parts. Although he initially referred to the falsetto as being in between the chest and the head register, he changed his terminology to avoid any misunderstanding. In his new model, he renamed the falsetto register to the middle register and declared it separate from the head voice:

> The registers are as follows: The chest, the medium, the head . . . The chest voice, which has much greater power of vibration than the medium, requires, accordingly, a more determined contraction of the glottis; and this contraction is

> most easily effected by the enunciation of the vowel E. The medium is generally the more veiled of the two, and requires a greater expenditure of air. These two registers, in their lower notes, set in vibration the entire length of the glottis; and, as we have before observed, the gradual ascending of the sounds in the vocal scale causes the cartilages to come more and more into contact, the vibration being effected at last by the tendons alone. By the latter the glottis forms in female voices, the notes called head register. (Garcia & Garcia, 1912, pp. 3-4)

This passage suggests that Garcia believed the medium register to be a coordinated mix of the chest and head registers. This is further supported by the writings of some more modern pedagogues such as Cornelius Reid. In the following passage, Reid defends the two-register model and explains the true nature of the middle register by quoting the renowned British surgeon, Sir Morell Mackenzie:

> The definition given by Mackenzie in the *National Encyclopedia* (1886), however, partially clarifies the issue. He states, "The former is termed in the Italian school the *voce di petto*, or chest register, and the latter the *voce di testa*, or head voice. To these the Italians add another which joins the two registers and which partakes of the character of both; it is named the *mezzo falso*, or middle falsetto." The *mezzo falso* is identical with the *voce di finte*, or "feigned" voice. (Reid, 1950, p. 69)

Thus, although there were proponents of the three-register model at the time, the third register sprung forth as a result of mixing the original two registers. It is then safe to conclude that the old Italian school highlighted two distinct vocal registers with the goal of mixing them in such a way as to render the transition imperceptible.

Although modern vocal pedagogy points to several diverging theories on vocal registers, the two-register theory is the best supported by recent scientific findings. As mentioned earlier in this document, there are two basic laryngeal modes of vibration. In the first configuration, called the modal voice or chest voice, the thyroarytenoid (TA) muscles actively contract during vibration, causing the folds to resist the longitudinal tension of the cricothyroid (CT) muscles. This results in all layers of the vocal folds entering into vibration. On the other hand, in the second configuration, often called the falsetto, the TA muscles relax, causing the CT muscles to stretch them further. This stretching results in only the outer mucosal and ligamentous edges of the vocal folds entering into vibration. These two modes of vibration, although not completely understood by teachers of the time, were certainly acknowledged as evidenced by the previous excerpts.

One of the main reasons underlying the advent of modern multiple register theories is the phenomenon of acoustic registers. Although there are only two distinct laryngeal registers as mentioned above, there are several resonance registers. These registers, unlike laryngeal registers, are not caused by changes in vibratory patterns of the vocal folds. Instead, they are brought about by the shifts in resonances of the vocal tract. Although these resonance transitions are perceived by the singer as lifts or even breaks, it must be highlighted that the mode of vocal fold vibration does not change.

Ear Training and Expressivity

Although the role of the ear is not given as much importance in the modern study of singing, the early teachers of the *bel canto* tradition certainly acknowledged the importance of the ear to the art of singing. A well trained ear was equally as important as the voice itself (Duey, 1980, p. 90). The early teachers recognized the fact that a good ear is a necessary requirement for good singing. In fact, the Spanish theorist Domenico Pietro Cerone, highlighting the control of the ear over the voice, claimed that an advanced singer sings more with the ear than with the mouth (Duey, 1980, pp. 90-91).

Despite this strong sentiment in favor of the ear, not much advice was offered early on regarding the cultivation of a good ear or the repair of a poor one. It was not until Mancini's treatise that more detailed instructions regarding the ear were provided. Mancini was of the opinion that singing out of tune was the worst fault and that singing with a nasal or throaty voice was to be preferred over out-of-tune singing.

> All the other defects can be corrected if not completely, at least to a point, of being noticeable to those only who are familiar with the art of singing. Singing out of pitch cannot be overcome nor disguised . . . The natural [dissonance] is a defect of Nature, who has not endowed certain youths with a sensitive ear for harmony; hence they are not able to distinguish between consonance and dissonance of tones. In such cases, correction is impossible. (Mancini & Buzzi, 1912, p. 62)

Although he stated that there are some defects of the ears that cannot be fixed, he offered advice to the teacher regarding their role in discerning the causes of a poor ear:

> It is the stern duty of the teacher to examine and find out what causes the dissonance. This is not difficult to do, but it requires wide experience. One must go very slowly and observe every detail in order to discern causes. The student should be tried at singing early in the morning before eating; and during the day, when the sky is cloudy and also when it is serene; when the air is placid and tranquil, on windy and stormy days; also soon after a full meal . . . The dissonance caused accidentally, can be corrected if the teacher is competent to discover and eliminate the cause. Among the temporary causes from which one sings out of pitch, is weakness of chest resulting from disease or temporary illness, or sometimes caused by indigestion, by eating too much or at no regular time or from other similar disorders . . . Now if the student sings out of pitch today and sings in pitch tomorrow and recognizes his mistake, then, there is no doubt but that the dissonance is accidental and can be corrected. (Mancini & Buzzi, 1912, pp. 63-64).

Mancini's advice, though not presented with scientific proof, was undoubtedly backed with years of experience and observations of his own students. Mancini did not provide reasons why some ears do not recognize pitch as well as others; however a modern scientific understanding of auditory mechanics, which was established earlier in this document, can assist in clarifying this.

The ear is tonotopic, as mentioned earlier, meaning that each pitch is mapped to a specific place in the cochlea. In essence, when a frequency of 440 Hz (which corresponds to the pitch A4) enters the cochlea, it excites a specific place on the membrane that is only sensitive to that particular frequency. This mapping of frequency to place, which begins in our cochlea, is preserved throughout the entire central auditory nervous system. In cases of amusia, or tone-deafness, the subject is unable to discriminate between two pitches. However, this is not due to the impairment of the pitch perception mechanisms described above at all; instead it is attributed to the failure to communicate these responses to higher brain centers. In an experiment by Peretz, Brattico, Jarvenpaa, and Tervaniemi (2009), it was found that individuals with amusia exhibited the same brain responses as normal individuals in response to mistuned notes. These results suggest that the distinguishing factor between amusic and normal brains is not the lack of ability to perceive small pitch differences, but the absence of awareness of the ability. The authors concluded that the weak connection between the auditory cortex and the inferior frontal gyrus in amusic brains is what prevents the conscious awareness of these pitch differences. In other words, their brain is able to detect these differences, but their mind remains unaware.

Just as important as the ear was to the *bel canto* tradition was the aesthetic expression of affect and emotion. After all, beauty is an integral part of beautiful singing. It is therefore necessary to briefly mention here the aesthetics of the *bel canto* tradition. Expressivity was undoubtedly of the highest priority to Caccini. In fact, he introduced his new style of monody with the primary goal of heightening the musical expression of the solo voice. In his treatise he emphasized the importance of affect:

> And yet I have never been content within ordinary boundaries accepted by others; rather have I always proceeded to seek out all the novelty I could, provided that the novelty be such as to facilitate attaining the goal of music, namely to give delight and to move the affect of the soul. (Caccini & Hitchcock, 1970, p. 49)

In the same treatise, he also described the various ornaments and their proper use to enhance the affect of the music. According to Caccini, these ornaments, including vocal crescendos and decrescendos, *esclamazioni*, tremolos and trills, and others, had the power to move others when used correctly. Perhaps the ornament that he describes with greatest enthusiasm is the *esclamazione*:

> Now, an *esclamazione* is really nothing but a certain strengthening of the relaxed voice; and the vocal crescendo in the soprano range, especially with falsetto, often becomes harsh and unbearable to the ear, as I myself have heard on many occasions. Without a doubt, therefore, as an affect more apt to move [the listener], a better result will be had from decrescendo on the attack than from a crescendo; for in the first manner—the [attack with a] crescendo—to make an *esclamazione* one must after relaxing the voice crescendo even more, and thus, I say, does it seem strained and coarse. But a wholly different result is obtained by [an initial] decrescendo, since at the point of relaxation giving it just a bit more spirit will make it ever more affective. (Caccini & Hitchcock, 1970, p. 49)

Caccini's style of music, with the aid of devices such as the *esclamazione*, took on a noble character which he referred to as *sprezzatura*. In essence, *sprezzatura* was characterized by a nonchalant rhythmic flexibility that allowed for a graceful flow of the words. To many, Caccini's style of singing evoked a performance aesthetic known as *meraviglia* (Stark, 1999, p. 160). This term, which literally translates to "the marvelous," was used to describe virtuoso singers who would elicit a sense of wonder, surprise, or the supernatural from the audience. This is particularly important because virtuosity and beauty together comprised the twin ideals of the *bel canto* tradition. These two elements were preserved and carried on throughout all the golden ages of singing. Thus, it is evident that Caccini's style of singing laid the foundation for *bel canto*.

A modern perspective of expression and emotion was laid out earlier in this document through the review of recent research in the field. Although many of the findings presented in that chapter are too scientifically advanced to be mentioned in the *bel canto* literature, there are some findings that provide support for certain *bel canto* ideals. For example, Timmers and Ashley (2007) found that ornamentation was a powerful tool for the expression of basic emotions through the variance of certain acoustical elements including loudness, duration, timing, complexity, and density. This is directly related to Caccini's philosophy of singing. The ornaments which he describes in his treatise are essentially stylistic variances of acoustic phenomenon (e.g., the *esclamazione* varies loudness over a period of time). Thus, this study provided some evidence for Caccini's view that ornaments, when used correctly, can indeed express emotion.

Further support of Caccini's view comes from Patrick Juslin's (2003b) GERMS model. This model, as described earlier, systematically divides expression of emotion into five distinct components. Two of these components are of particular interest: movement patterns and stylistic novelty. Movement patterns refers to the biologic movement patterns characteristic of life and nature. Such patterns, when present in music, enhance expression. This concept resonates with many of Caccini's ornaments, which also exhibit these patterns. Stylistic novelty, which refers to the performer's deviations from stylistic expectations in order to add an element of unpredictability, is also closely connected to Caccini's style. The *sprezzatura*, or rhythmic flexibility, present within his style created unpredictable yet aesthetically pleasing nonchalant variations of rhythmic flow that elicited a sense of *meraviglia* or wonder in the listener.

Legacy

Although *bel canto* was primarily an empirical tradition passed down through oral and aural tradition and was not founded upon scientific or acoustical knowledge, the present document has aimed to show that many of its precepts are undeniably grounded in scientific principles. It is extremely valuable for teachers of singing to understand these underlying principles; by doing so, they will be well equipped to satisfy the queries of inquisitive young singers and explain why the age-old principles do indeed work. On the other hand, without this understanding, it will be difficult to discern the reasons underlying the traditions and practices. One of the main insights this project has brought to light is the fact that although science and tradition are often viewed as being at odds with one another, they are not mutually exclusive of one another. In fact, many of the *bel canto* teachers attempted to explain their teaching philosophies systematically and scientifically, but for lack of sufficient scientific research, could not do so entirely.

It is certainly not the aim of this document to persuade all voice teachers to adopt a scientific method of instruction; in fact, this method of instruction may actually do more harm than good for some students. Nevertheless, the general knowledge of the scientific principles underlying each practice is essential for students in order to keep on track amidst the varying perspectives in vocal pedagogy. Even within the *bel canto* tradition it was quite apparent that there were many differing views. However, with a deeper understanding of the underlying concepts, it becomes more and more evident that many of the seemingly opposing views are actually just emphasizing different aspects of the same principle. The Lamperti school, for example, taught in a very different manner than the Garcia school, but upon closer inspection, it was evident that their ideologies were not mutually exclusive.

This idea can be better expressed through the analogy of the old folktale of the blind men and the elephant. According to this story, six blind men who have never seen an elephant before in their lives decide to visit one and describe it based on touch and feel. However, each man touches a different part of the elephant and consequently describes it focusing on that aspect alone. The first man, touching the tail, describes the elephant as a rope. The second man describes it as a tree based on his experience of feeling its sturdy legs. The third, feeling the sharp tusks, compares it to a spear. The fourth describes the elephant as a snake after touching its curved trunk. The fifth, feeling its ears, describes it as a large hand fan. Finally, the sixth describes the elephant as a large wall based on feeling its side. This story illustrates that while all of the men certainly did describe the individual parts of the elephant quite well, their descriptions were at odds with each other. However a deeper understanding regarding the interrelatedness between these parts may have provided the men with more insight into the appearance of the elephant.

In the same way, it is important to keep in mind that while there are many different opinions and methods regarding singing, they may simply be focusing on different aspects of each concept. Thus, in order to obtain a more comprehensive perspective, it is important to obtain a deeper understanding of the scientific principles linking these puzzle pieces together.

In closing, it must be reiterated that science and tradition should not be seen as opposite ends of the spectrum; instead they should function together to enhance vocal instruction and create new golden ages of singing. However, in order to accomplish this noble goal, further bridging of the schism between tradition and science is necessary. Only when there is a union between science and tradition can we truly make progress in the field of vocal instruction, continuing to teach and inspire students in the art of beautiful singing.

References

Adachi, M., & Trehub, S. E. (2000). Decoding the expressive intentions in children's songs. *Music Perception: An Interdisciplinary Journal, 18*, 213-224.

Appleman, R. (1974). *The science of vocal pedagogy: Theory and application*. Bloomington, IN: Indiana UP.

Barbier, P. (1996). *The world of the castrati: The history of an extraordinary operatic phenomenon*. London: Souvenir.

Basic structure of the respiratory system. (n.d.) Retrieved from http://www.wpclipart.com/medical/anatomy/lungs/respiratory_system.png.html

Berkeley, G. (1710/1957). *A treatise concerning the principles of human knowledge.* New York, NY: Liberal Arts Press.

Caccini, G., & Hitchcock, H. W. (1970). *Le nuove musiche*. Madison, WI: A-R Editions.

Celletti, R. (1991). *A history of bel canto*. Oxford [England]: Clarendon Press.

Components of the ear. (n.d.) Retrieved from http://www.wpclipart.com/medical/anatomy/ear_diagram.png.html

Correspondence. (1884, April 5). Notes and Observations. *Scientific American,* 218.

Duey, P. A. (1951). *Bel canto in its golden age: A study of its teaching concepts*. New York, NY: King's Crown Press.

Fant, G. (1960). *Acoustic theory of speech production: With calculations based on x-ray studies of Russian articulations*. 's-Gravenhage: Mouton.

Ferrand, C. T. (2007). *Speech science: An integrated approach to theory and clinical practice*. Boston, MA: Pearson/Allyn and Bacon.

Franca, I. (1959). *Manual of bel canto*. New York, NY: Coward-McCann.

Garcia, M., & Garcia, A. (1924). *Garcia's treatise on the art of singing. A compendious method of instruction with examples for the cultivation of the voice*. London: Leonard.

Gray, H., Lewis, W. H., & Gray, H. (1918). *Anatomy of the human body.* Philadelphia, PA: Lea & Febiger.

Hirano, M., & Bless, D. M. (1993). *Videostroboscopic examination of the larynx*. San Diego, CA: Singular Publishing Group.

Hodges, D. A., & Sebald, D. C. (2011). *Music in the human experience: An introduction to music psychology*. New York, NY: Routledge.

Hoit, J. D., Jenks, C. L., Watson, P. J., & Cleveland, T. F. (1996). Respiratory function during speaking and singing in professional country singers. *Journal of Voice, 10*(1), 39-49.

Juslin, P. N. (1997). Emotional communication in music performance: A functionalist perspective and some data. *Music Perception: An Interdisciplinary Journal, 14*, 383-418.

Juslin, P. N. (2003a). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin, 129*, 770-814.

Juslin, P. N. (2003b). Five facets of musical expression: A psychologist's perspective on music performance. *Psychology of Music, 31*, 273-302.

Konečni, V. J. (2008). Does music induce emotion? A theoretical and methodological analysis. *Psychology of Aesthetics, Creativity, and the Arts, 2*, 115-129.

Kulkarni, N. V. (2012). *Clinical anatomy: A problem solving approach*. New Delhi: Jaypee Brothers Medical Publishers.

Lamperti, F. (1864). *Guida teorico-pratica-elementaire per lo studio del canto*. Milano: Tito di Gio. Ricordi.

Lamperti, F., & Griffith, J. C. (1890). *A treatise on the art of singing*. New York, NY: G. Schirmer.

Lamperti, G. B., & Brown, W. E. (1957). *Vocal wisdom: Maxims of Giovanni Battista Lamperti*. Boston, MA: Crescendo Publishing.

LeBlanc, A., & Sherrill, C. (1986). Effect of vocal vibrato and performer's sex on children's music preference. *Journal of Research in Music Education, 34*(4), 222-237.

Liberman, M. C., & Gao, J. (2002). Prestin is required for electromotility of the outer hair cell and for the cochlear amplifier. *Nature, 419,* 6904.

Livingstone, S. R., Thompson, W. F., & Russo, F. A. (2009). Facial expressions and emotional singing: A study of perception and production with motion. *Music Perception: An Interdisciplinary Journal, 26*(3), 475-488.

Lunn, C. (1862, January). The two schools: Garcia v. Lamperti [Letter to the editor]. *The Musical Standard: A Newspaper for Musicians, Professional and Amateur*. London: Reeves and Turner.

MacCormac, W., & Makins, G. H. (1881). *Transactions of the International Medical Congress: Seventh session held in London, August 2d to 9th, 1881*. London: J.W. Kolckmann.

Mancini, G., & Buzzi, P. (1912). *Practical reflections on the figurative art of singing*. Boston, MA: R.G. Badger.

Martin, F. N., & Clark, J. G. (2009). *Introduction to audiology*. Boston, MA: Pearson.

McKinney, J. C. (2005). *The diagnosis & correction of vocal faults: A manual for teachers of singing and for choir directors.* [Reissue]. Long Grove, IL: Waveland Press.

Miller, R. (1996). *The structure of singing: System and art in vocal technique*. New York, NY: Schirmer Books.

Müller, J., & Baly, W. (1848). *The physiology of the senses, voice, and muscular motion, with the mental faculties*. London: Taylor, Walton & Maberly.

Peretz, I., Brattico, E., Järvenpää, M., & Tervaniemi, M. (2009). The amusic brain: In tune, out of key, and unaware. *Brain: A Journal of Neurology, 132,* 1277-1286.

Pilotti, K. (2009). *The road to bel canto* (Masters Thesis). Retrieved from http://oru.diva-portal.org/smash/get/diva2:291620/FULLTEXT02.pdf

Pleasants, H. (1966). *The great singers: From the dawn of opera to our own time*. New York, NY: Simon and Schuster.

Reid, C. L. (1950). *Bel canto: Principles and practices*. New York, NY: Coleman-Ross.

Rosen, S., & Howell, P. (1991). *Signals and systems for speech and hearing*. London: Academic Press.

Sataloff, R. T. (2005). *Voice science*. San Diego, CA: Plural Publishing.

Sataloff, R. T. (2013). *Atlas of laryngoscopy*. San Diego, CA: Plural Publishing.

Scherer, K. R. (1986). Vocal affect expression: A review and a model for future research. *Psychological Bulletin, 99*, 143-165.

Scherer, K. R. (1987). Toward a dynamic theory of emotion: The component process model of affective states. *Geneva Studies in Emotion and Communication*, 1-98.

Scherer, K. R. (1995). Expression of emotion in voice and music. *Journal of Voice, 9*, 235-248.

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication, 40*, 227-256.

Seikel, J. A., King, D. W., & Drumright, D. G. (1997). *Anatomy and physiology for speech and language*. San Diego, CA: Singular Publishing Group.

Stark, J. A. (1999). *Bel canto: A history of vocal pedagogy*. Toronto: University of Toronto Press.

Timmers, R., & Ashley, R. (2007). Emotional ornamentation in performances of a Handel sonata. *Music Perception: An Interdisciplinary Journal, 25*, 117-134.

Titze, I. R. (2008). Nonlinear source-filter coupling in phonation: Theory. *The Journal of the Acoustical Society of America, 123*(5), 2733-2749.

Toft, R. (2013). *Bel canto: A performer's guide*. New York, NY: Oxford University Press.

Tosi, P. F., & Galliard, J. E. (1968). *Observations on the florid song*. New York, NY: Johnson Reprint.

Vennard, W. (1967). *Singing: The mechanism and the technic*. New York, NY: C. Fischer.

Watson, P. J., & Hixon, T. J. (1985). Respiratory kinematics in classical (opera) singers. *Journal of Speech and Hearing Research, 28*(1), 104-122.

Willis, N. P. (1851). *Memoranda of the life of Jenny Lind*. Philadelphia, PA: Robert E. Peterson.

Zañartu, M., Mehta, D. D., Ho, J. C., Wodicka, G. R., & Hillman, R. E. (2011). Observation and analysis of in vivo vocal fold tissue instabilities produced by nonlinear source-filter coupling: A case study. *The Journal of the Acoustical Society of America, 129*(1), 326-339.